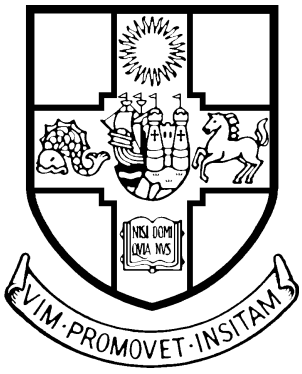
Tuneable Synthetic Genetic Devices

Stay Tuned

By

VITTORIO BARTOLI



Department of Engineering Mathematics
UNIVERSITY OF BRISTOL

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY in the Faculty of Engineering.

MARCH 2020

Word count: 31,282

Abstract

Synthetic genetic circuits are gene regulatory networks used to engineer biological systems to carry out useful functions. Moving circuits between environments or host cells alters their function in ways we can't predict, leading to circuit failures that can only be fixed by laboriously rebuilding them. Here, we propose a novel regulatory motif that controls transcription and translation of a gene. We implemented the motif to design a device - a tuneable expression system (TES) that used a riboregulator to activate translation of a gene in response to cognate small RNAs (sRNA). Transcription of the gene and sRNA were independently regulated by two sensors that were activated in response to different inducers, allowing us to dynamically tune the device's response function. The TES's outputs at high and low inputs could be shifted 4.5 and 28-fold, respectively. We tested the TES in a range of glucose concentrations to emulate how these conditions would affect device performance in industry and found that it produced protein >2-times faster in higher glucose concentrations. We showed that the TES can be regulated so protein production rate remains constant in different glucose concentrations. We then used the TES to build a tuneable repressor protein based NOT gate, whose transition between on and off states can be tuned over a >6-fold range. However, in all devices, tuning the device reduced its fold-change and separation between populations of cells with high and low inputs. Using deterministic and thermodynamic models we found we could improve the device's performance by increasing the rate of sRNA transcription and removing a self-cleaving ribozyme insulator that interfered with riboregulator function. Circuits built using the TES could be tuned and fixed dynamically, removing the need to reassemble them from new parts and accelerating genetic circuit development. Furthermore, TESs provide the basis for novel adaptive circuits, systems that self-regulate their behaviour to be optimal in all condition.

Dedication and acknowledgements

I dedicate the energy used in my research and writing to all those suffering as a result of climate change.

- Thanks to Tom and Mario. The time and patience you've put into my cultivation is deeply appreciated.
- Thanks to my pals in life sciences, everyone in lab 140 and office 107.
- Thanks Danny, Marco, Veronica and Matt, you've brought me much joy.
- Thanks to my fellow rebels for bringing color and hope into my life.
- Thanks to all my family for their constant love and support.
- Thanks to the beez for their absurdities.
- Thanks to the cook club for the suggestions on how to use GFP in new and exciting ways.
- Thanks to all those in BrisSynBio who I've crossed paths with, especially Kathleen and Viv.
- Thank you Andy and Lorena for your help with flow cytometry.
- Thanks to Eleanor for showing me how to iron out the creases in my mind, body and soul.

This body of work was supported by the EPSRC/BBSRC Centre for Doctoral Training in Synthetic 589 Biology grant EP/L016494/1 and BrisSynBio, a BBSRC/EPSRC Synthetic Biology Research Centre grant 588 BB/L01386X/1.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE: .....

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

Synthetic biologists re-write the code of life, DNA, to engineer living systems with their own desired behaviour. Examples include engineered microbes that can: produce sustainable palm oil [15], convert agricultural waste into building materials [12] and brew beer that doesn't need hops adding to it [117]. This has been made possible by advances over the past 50 years that have identified the function of many genes and regulatory sequences encoded in DNA. Synthetic biologists exploit this knowledge to build new living systems, preferably in a systematic and rational way.

1.1 Engineering principles applied to biology

To more predictably modify life, synthetic biologists often apply principles developed by engineers to simplify the complex processes that occur in living systems. This is achieved by focusing on the key features for the engineering task at hand. These simplified representations of core biological processes are combined at different scales to build systems of increasing complexity (**Figure 1.1**). At the smallest scale, DNA sequences are thought of as interchangeable parts that carry out specific functions. These DNA parts are combined to build devices that carry out more complex functions. Parts and devices can then be combined to build larger genetic circuits. Like electronic circuits, these sequences of DNA detect an input, perform computations to process this information, then actuate a desired response. To engineer DNA sequences in this way synthetic biologists employ four core engineering principles:

1. **Abstraction** – engineers use simplified representations to describe processes in biological systems in terms of the functions they perform that are of interest. It's used across all different scales of system in biology. We describe a sequence of DNA that carries out a specific function as a part (**Figure 1.1**), for example, transcriptional promoters are sequences of DNA that define where an enzyme, RNA polymerase (RNAP), starts to

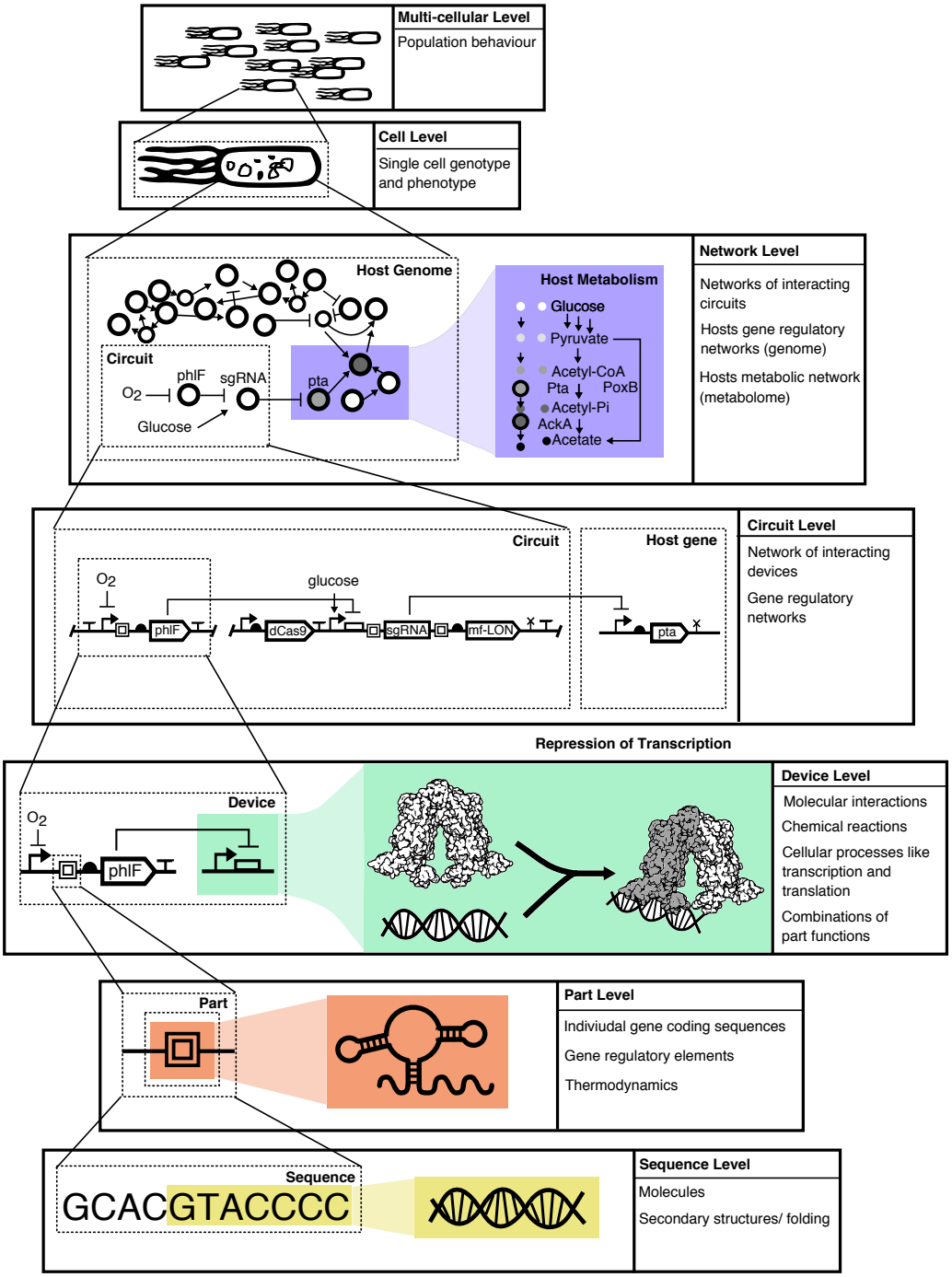


Figure 1.1: Synthetic biology hierarchy A schematic diagram demonstrates the hierarchy of abstractions used in synthetic biology. As an example, we use a genetic circuit that manipulates the metabolism of a host cell [96]. Systems increase in complexity from the smaller scale systems (bottom) to larger scale systems (top). Coloured regions demonstrate abstraction. In yellow, DNA molecules are represented as sequences of letters. In orange, the function of a genetic part is determined by RNA folding. In green, transcription of a gene is prevented by a protein that binds to DNA, preventing it being transcribed by RNA polymerase (RNAP). In purple, the genetic circuit produces less acetate when glucose and oxygen (O_2) are present.

transcribe a gene. When using a promoter, synthetic biologists are not normally concerned with the bases that make it's sequence, they just place the abstracted part upstream of a gene of interest.

2. **Composability** – engineers intentionally build components that can be used interchangeably, so their functions are maintained in any context. Composability makes it simpler to build large systems from smaller, modular components (modules). Modules carry out the same function when they are used on their own or in a system, allowing us to predict how a large system composed of modules will behave based on how the modules behave on their own. This makes it simpler to build complex living systems.

3. **Decoupling** – components are designed so that their activity does not alter the behaviour of other components in a system. In biological systems, components are made up of biochemical molecules that constantly diffuse around a cell. Single molecules that are constituents of one module can affect several others. For example, if a protein in a module binds not only to a specific sequence of DNA we desire, but to other similar sequences in different modules, it will affect the behaviour of the other modules, changing how they behave. This effect leads to unpredictable changes in the system's behaviour, so to predictably build larger systems, modules must be decoupled.

4. **Standardization** – components are built and tested following a set of rules that allow components to be easily connected together. Using a standardized signal carrier as the input and output to modules, the output from one module can be used as the input to another [19], making them easier to connect. For example, by standardizing all devices to use transcriptional promoters as device inputs and outputs, RNAP flux can be used as a common signal carrier across devices [13, 19, 100]. Other signal carriers include the concentration of the active form of a protein or ribosome flux on an mRNA [90].

1.2 The design, build, test, learn cycle

Successfully engineering complex systems with specified functionalities is difficult to accomplish on a first attempt so a systems development life cycle approach is used. A product is designed, built and tested in comparison to the initial specification. If the product doesn't meet the specification, further cycles are taken until a functional product is conceived. Engineers learn from failed designs to improve subsequent designs until one is found that fully meets the specification. Synthetic biologists have adapted this framework to engineer biological systems and manage the complexity and large uncertainty associated with the task (**Figure 1.2**).

To create a genetic part, device or circuit, the first step is to develop a design. In the design stage, we aim to predict the function of a specific DNA sequence. Knowledge from previous

experiments and mathematical modelling is used to predict how components will behave when combined (**Figure 1.2**).

The build stage of the cycle is generally the most time consuming and expensive. The specific sequence of DNA encoding a design needs to be accurately constructed into DNA. For the construct to be tested it needs to be built into a vector which enables the function encoded in the sequence to be executed.

Once constructs have been built and sequence verified they need to be tested to see if they carry out the function they were built for. The way devices are tested depends on their end use. If a system's performance does not meet specification it needs debugging. This involves identifying which parts or devices aren't performing as expected. We then learn from this and modify the design on the next iteration of the DBTL cycle.

In the learn step, we analyse the results from testing and generate ideas for improving the design. The DBTL cycle has been shown to be an effective way of developing and optimizing engineered genetic systems [20]. However, every extra iteration of the cycle costs time and money. Therefore, reducing the number of times a construct needs to be redesigned and built is highly beneficial.

Another challenge is that genetic circuits are tested in controlled lab based environments, irrespective of the environment that the circuit is intended to be used in. For genetic systems to be used in a real-world context, they need to have robust functions and be able to adapt

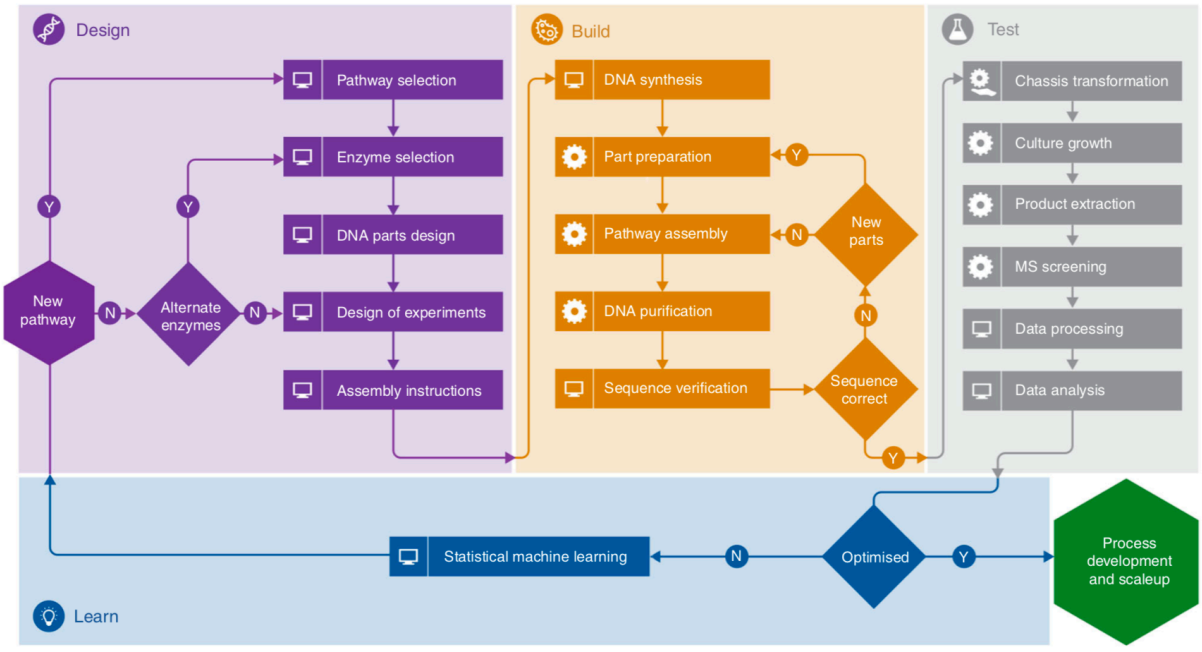


Figure 1.2: The design, build, test, learn cycle. Flow diagram of the design-build-test-learn cycle used to engineer microbes to produce valuable chemicals (reprinted from *Commun. Biol.*, 1, Carbonell *et al.*, An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals, 66, (2018) [20].)

to environmental perturbations. However, circuits are often tested in carefully controlled lab conditions and break when deployed to complex, real-world environments [21, 52, 95].

1.3 Motivation

The time taken to create devices that function precisely as desired is often prohibitively long. A specific example of this is the creation of new metabolic pathways in industrial biotechnology. In this context, genetic circuits are used to regulate the rate that key enzymes involved in metabolic pathways are produced. Correct regulation of these enzymes allows for the efficient production of valuable chemicals. However, the DBTL cycle used to create such pathways is time consuming and expensive. It is estimated that it takes 50 to 300 person years and millions of pounds to develop a viable production strain [80]. For synthetic biology to have the greatest impact on society, we need ways to rapidly scale developments made in the lab to industrial contexts. We can do this by reducing the time taken for a DBTL cycle to be carried out and the number of cycles needed to develop a functioning device or circuit.

The motivation of this project was to develop tools and methodologies that make it faster and cheaper for synthetic biologists to engineer working genetic circuits. More specifically, we aimed to build genetic devices whose behaviours can be dynamically modified after their assembly in a circuit. This innovation would cut the cost and time taken to build genetic systems by removing the need to reassemble the circuit for each DBTL cycle. If the circuit doesn't work as intended, we can simply tune the devices until the circuit is fixed, with no reassembly of the encoding DNA required. This approach focuses on broadening a circuit's functionalities encoded in a single design-space – the subset of DNA sequences that could be used to encode a device or circuit, to improve its performance. However, changing the environment a genetic system is tested in also alters its performance. Therefore, a genetic component's test-space – the set of environmental conditions that circuits are tested in, is just as important. A second aim was to test genetic circuits in environments that emulate the conditions they will be exposed to in the real world, allowing us to uncover otherwise unforeseeable design flaws. Testing genetic circuits in appropriate test-spaces could reduce the time taken for their use to be scaled from the lab into real-world applications.

1.4 Thesis overview

After presenting some background in Chapter 2, in Chapter 3 we describe the materials and methods used to undertake this research. In Chapter 4 we present the idea of a device capable of tuning the rate that a gene is expressed by providing us with a way to independently regulate rates of transcription and translation. We call this device a tuneable expression system (TES) and use mathematical modelling to refine its design. We go on to validate the design by building a plasmid implementation and testing this *in-vivo*. From experimental data, we show that the

device doesn't perform well and so identify issues with the initial design that could potentially hinder its performance.

In Chapter 5, we use a combination of deterministic, stochastic and thermodynamic modelling to test these hypotheses *in-silico* and based on these results, design new variants of the device. We then compare the performance of these optimised devices to the original design.

Finally, in Chapter 6, we use the TES to create a tuneable NOT gate that could be used to implement biological computations. We build and test four variants and compare their performance *in-vivo*. We then test how glucose concentration affects the performance of the TES.

Excluding Sections 4.5 and 6.4, the contents of this thesis are published in: Vittorio Bartoli, Grace A. Meaker, Mario di Bernardo and Thomas E. Gorochowski, "Tunable genetic devices through simultaneous control of transcription and translation", *Nature Communications*, 2020 [4].

CHAPTER 2

Background

Humans first used biology as technology more than 23,000 years ago, when they first cultivated plants for sustenance [123]. This innovation was the driving force behind the transition of societies from being nomadic to being settled, it drastically changed how humans live and literally changed the face of the Earth [107]. New crops were engineered through the slow process of selection. Thousands of years after the advent of agriculture came another revolution in biotechnology. The first evidence of fermentation arose at the same time as the first signs of civilization around 7000 years ago [91]. Fermenting sugars to produce alcohol allowed our ancestors to preserve food and provided a valuable product that was used as payment for services, allowing leaders to leverage people power and pursue more ambitious urban projects, leading to the birth of civilization itself [91].

2.1 The development of biotechnology

The next big step in shaping biotechnology as we know it came in the late 1500s when microscopes were invented [131]. Robert Hooke and Antoni van Leeuwenhoek used microscopes to see microorganisms (microbes) for the first time. 150 years later, microscopy helped us understand how microbes play an important role in causing diseases and transforming chemicals [47]. The next major innovations in biotechnology came thanks again to alcohol. Whilst researching the production of wine and beer in 1857, Luis Pasteur found that fermentation was caused by microbes and not the spontaneous decomposition of chemicals, as was previously believed [131]. These discoveries led to the foundation of modern day biotechnology.

The next big wave of innovation in biotechnology came with the discovery of DNA in the 1950s. DNA was shown to contain information that cells used to produce proteins. In 1978, Warner Arber, Hamilton Smith and Daniel Nathans won the Nobel prize for their work on endonucleases, bacterial enzymes that cut the DNA of invading viruses [131]. The discovery of endonucleases,

along with other innovations at this time, enabled scientists to "edit" DNA. They could "cut" out and build new fragments of DNA designed to carry out novel functions and place them in living organisms. The first industrial use of this technology was when scientists engineered *Escherichia coli* to produce insulin in the 1970s [131]. This was the first protein produced commercially by a microbe engineered with this technology.

Biotechnology has always been closely related to economics because funding research and development is expensive [131]. Funding comes from private or public investors who expect innovations to directly or indirectly benefit the economy, so it's crucial that biotechnological innovations lead to profitable enterprises. The most recent set of innovations in biotechnology led to the advent of synthetic biology. Cheaper DNA sequencing and synthesis has enabled scientists to build and test large numbers of genetic designs and to better understand the rules for reprogramming life. This has led to a more rational approach to engineering biology, which has the potential to change how we produce many of the chemicals and materials we rely on as a society.

2.2 Money steers innovation

Microbes can be engineered to convert waste products into valuable chemicals [82]. For example, waste products from agriculture can be put through industrial processes that use engineered microbes to convert them into fuels, pharmaceuticals and materials [82, 101]. For these processes to be commercialized they need to be profitable. A fast way of measuring if a process would be profitable is to subtract the cost of raw materials (also called feedstock) from the expected revenues, a figure known as the gross margin [129]. The cost of raw materials is usually 80–90% of the cost of production, so the gross margin gives an accurate estimation of how profitable a process will be. A low gross margin suggests that a process is not commercially viable.

To make a process more profitable, two measures of its performance can be improved. First, the gross margin depends on how much of a raw material can be converted into a product, a value summarised in the yield, which is the percentage of a raw material converted into a product. Synthetic biology lets us engineer microbes to produce higher yields of a product. By increasing the yield, we generate more product from the same amount of feedstock, increasing the gross margin, making a process more profitable and increasing the chances of it being commercialized.

A second way of increasing the profitability of a process is to increase how much of a product is produced in a fixed amount of time – the process's productivity. Productivity is related to the scale of production; the volume of chemical being used and how rapidly microbes convert these into a desired product. If we increase scale of production, we can increase how much product we generate in the same length of time and therefore revenue, however, increasing the scale of production increases the size or amount of equipment needed to carry out a process. This increases the initial capital needed. If the microbes used in a process are more productive, the

same volume of product can be produced in a smaller plant, reducing building costs. Therefore, engineering productive microbes reduces a process's capital cost, making investors more likely to invest.

In the past, productivity, yield and process economics were improved by selectively growing microbes that produced the best beers, wines, breads and cheeses. Since the 1970s metabolic engineers have used faster, more effective strategies to engineer microbes for industrial processes. Valuable chemicals are often found in nature, produced by either plants or microbes [104]. However, these organisms are not always well suited to growth in an industrial plant.

One strategy metabolic engineers use to engineer production strains is to identify genes that code for sets of enzymes that produce valuable products [101, 104]. They place these genes in microbial strains that when grown at industrial scales, are able to produce the desired product at a higher yield and with better productivity than the native strains in which the pathway was found [80, 101, 104]. Genes that occur naturally in the metabolism of engineered strains produce enzymes that convert feedstock into worthless waste products. Thus, to further improve production strains, metabolic engineers turn off the expression of these genes to channel resources to those genes that produce the desired products [80]. These innovations have improved the yield and productivity of numerous processes currently being used at a commercial scale [101]. However, engineering strains in this way is time consuming and expensive.

For example, using this approach to develop a viable product requires 50–300 person years and several million's of pounds of investment [80]. This is where synthetic biology can help. New methods to synthesise large DNA fragments open up more rational approaches to engineer microbes for industry.

2.3 Improving bioproduction with synthetic biology

Cells produce proteins via the multi-step process of gene expression. Information encoded in DNA is first converted into RNA by RNA polymerase (RNAP) during a step called transcription. This RNA is then bound by ribosomes that translate it into a sequence of amino acids that fold to produce a protein. Controlling the rates of transcription and translation is crucial to cellular functions and synthetic biologists have developed ways of modifying these using biological parts. By regulating rates in new ways, we can engineer cells to produce proteins when they detect specific changes in their environments and thus, more effectively produce a product in less time.

A clear example of this approach is by Zhang et al. who improved the production of a biodiesel – fatty acid ethyl ester (FAEE) [142]. First they engineered a biosensor based on a naturally occurring promoter, P_{modB}, that increased the rate of transcription of genes when fatty acids were present in the cells. They used this sensor to regulate the transcription of four enzymes: pyruvate decarboxylase (*pdc*), alcohol dehydrogenase B (*adhB*), acyltransferase A (*atfA*) and fatty acyl-CoA synthetase (*fadD*). These enzymes work together to produce FAEE from fatty acids. By

only producing the enzymes when fatty acid was present, energy and nutrients were not wasted on enzyme synthesis until there was sufficient substrate for them to process. Furthermore, FAEE is made from ethanol and fatty acyl-CoA. Whilst a necessary intermediate in FAEE production, ethanol damages cells, which decreases the amount of FAEE that they produce [125]. In this system, ethanol is produced by enzyme *pdc*, which is only produced when fatty acid concentration is high enough to be converted into the intermediate fatty acyl-CoA, which then reacts with and sequesters ethanol. This reduces the toxic effect of the ethanol when producing FAEE. By only producing ethanol when it is needed, cells can grow to be healthy and produce FAEE more effectively, a strategy that resulted in a three–fold increase in yield, from 9% to 28% [142].

This approach can be taken even further. More recently Moser et al. [96] use a synthetic biology approach to engineer a strain of *E. coli* that produces less waste by-product during fermentation. The central metabolism of *E. coli* produces acetone from glucose [96]. In metabolic pathways where acetone is not converted into a useful product, less glucose is available to make the desired product, reducing yield. To overcome this problem, Moser et al. built a genetic circuit that turns off production of the enzyme that catalyses acetate formation when the concentrations of glucose and acetate are both high. To do this, they built a genetic circuit to process two input signals: the concentrations of glucose and acetate. When both acetate and glucose are abundant in the cell, the circuit turns off transcription of the enzyme that produces acetate. This causes less glucose to be converted into acetate, leaving more available to be converted into desired products.

Clearly, using genetic devices and circuits to control rates of protein production is valuable for engineering microbes to produce chemicals more effectively. It differs from traditional metabolic engineering approaches where whole sets of genes are introduced and merely overproduced in cells, in a non-optimal and uncontrolled manner.

2.4 Engineering genetic circuits

Genetic parts are interchangeable sequences of DNA that regulate gene expression. Many parts can be connected together to create devices that perform specific functions. Similarly, by connecting many devices together their functions can be combined to build more complicated gene regulatory networks, or genetic circuits. Genetic circuits process information in cells, detecting an input and generating a required output based on some computation. Circuits are often hierarchical in structure, requiring many to be connected together to produce more complex functionalities. To make it easier to connect genetic devices and circuits, it helps if their inputs and outputs use a standard type of signal. These are often core biological processes. Examples include: the rate that RNA polymerase is recruited by transcriptional promoters (RNAP flux) [19], the concentration of a protein and the initiation rate of ribosomes at a ribosome binding site (RBS) [90]. A genetic device can be characterised by the relationship between its input and output signals, which a device's response function captures at steady-state. Response functions

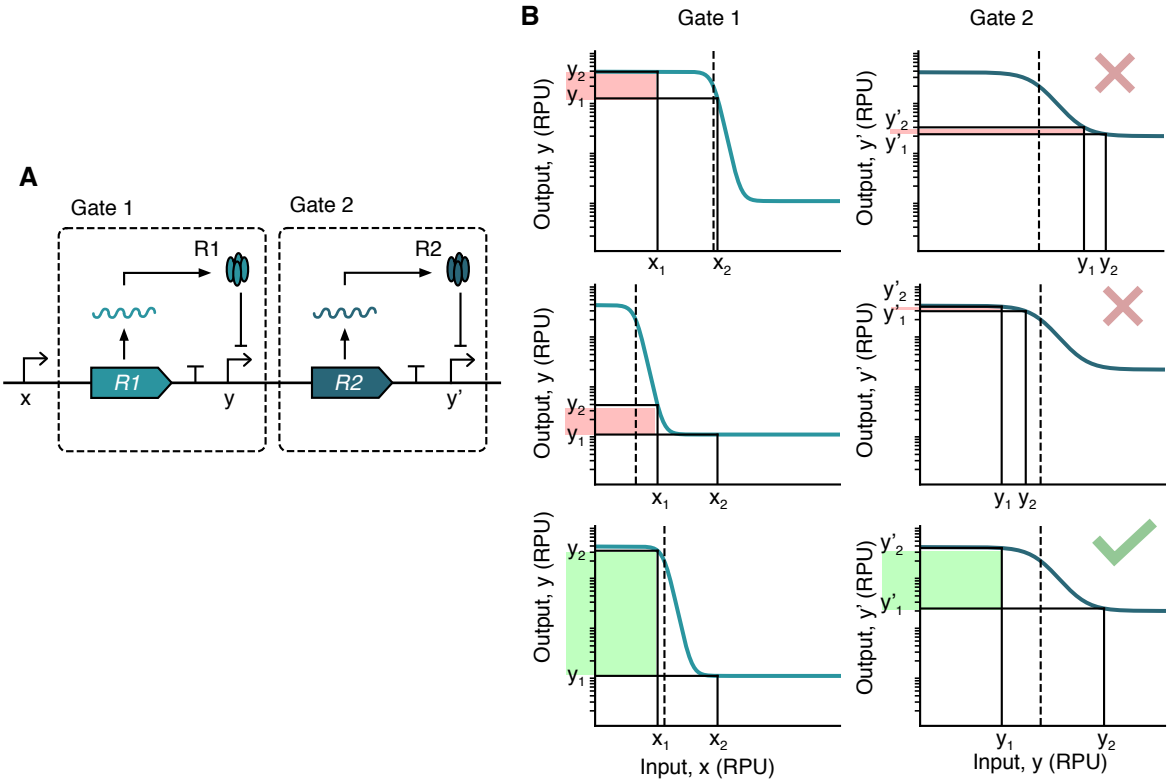


Figure 2.1: Illustration of mismatched genetic NOT gates. (**A**) Schematic of two NOT gates connected in series. x, is the activity of the input promoter to gate 1. $R1$ and $R2$ are repressor proteins that repress promoter activities y (which is the output from gate 1 and input to gate 2) and y' (which is the output from gate 2), respectively. (**B**) Response functions of gates 1 and 2 were modelled with reverse hill functions. The response function of gate 2 is fixed, whilst, from top to bottom, the K-value of gate 1 was simulated at 0.08, 0.001 and 0.005 RPU. High and low inputs to gate 1, x_1 and x_2, respectively, are the same in all simulations.

have become central to the rational design of complex circuits in living cells [100].

To connect two devices, their response functions need to be compatible, or matched, i.e. the range of outputs from the first device must span the necessary range of inputs to the second. Matching components is crucial when circuits are made from devices that rely on switching behaviours to carry out computations. For a device to exhibit switching behaviour, its outputs must switch between two states (significantly different high and low outputs) when its input is either high or low. If a change in the input signals doesn't cause the device to switch, the signal can't propagate through the circuit, causing it to break. The shape of a response function dictates whether the difference in input signals is large enough to shift the output signal between on and off states (**Figure 2.1**). By modifying the shape of a response function, the functions of two devices can be made to match. However, if the shape of matched response functions change when a circuit is being used, it can lead to the circuit failing [13, 130].

A few features of a device's response function indicate how well it will perform. To gauge

whether the outputs from a device will be sufficiently different to trigger switching in a down stream device, the ratio between the devices output in high and low states, also known as its fold–change is used. A device with a higher fold-change is more likely to trigger switching in a downstream device. However, caution must be taken in measuring a devices performance by fold–change alone. If the lowest output from a device is very low, it could result in a device having a large fold-change, even though the outputs do not span a large range. To account for this, the difference in a device's outputs in on and off states are reported. This measure of performance is the device's dynamic range.

Changes in environmental conditions such as temperature and the concentration of nutrients cells are grown in can all alter the behaviour of parts and devices [52, 95]. In some cases these changes are so great that response functions deviate from their required shape, causing entire circuits to break. Synthetic biologists claim that genetic circuits can be used in industrial biotechnology to coordinate gene expression and optimize processes. However, when cells are grown in industrial bioreactors, conditions such as temperature and concentrations of nutrients vary significantly over time and location in the reactor. Using cells in these highly variable environments will change the way that parts designed to be used under static conditions behave. If we are to create robust genetic systems, it is important that we can tune the function of the parts it consists of. One way to do this is by controlling the transcriptional and translational processes used by genetic parts and devices.

2.5 Tuning the rate of transcription

Transcription initiation is a complex process where the core RNAP enzyme interacts with a sigma factor to form a holoenzyme. The sigma factor portion of the holoenzyme recognizes a recognition site within a promoter sequence, which the holoenzyme then binds to [14, 56, 134]. The holoenzyme melts the DNA duplex around the transcription start site forming a transcription bubble; a process that yields a stable open promoter complex that marks the initiation of RNA synthesis [64]. Once the RNA chain reaches 8–11 nucleotides (nt), the transcription process transitions to the next step, productive elongation [10, 114]. There are many ways that this process can be influenced to allow for the rate of transcription to be controlled. The most common approaches are to use protein or RNA regulators.

2.5.1 Tuning rate of transcription with proteins

Proteins known as transcription factors (TF) are able to control transcription by altering the rate that RNAP binds promoter sequences. Repressors are TFs that bind an operator sequence. By placing an operator near or in a promoter the physical presence of the repressors blocks RNAPs from binding to the DNA, reducing the transcription initiation rate. Ideally, the repressor binds specifically to its cognate operator site. However, off-target binding can sometimes occur, which

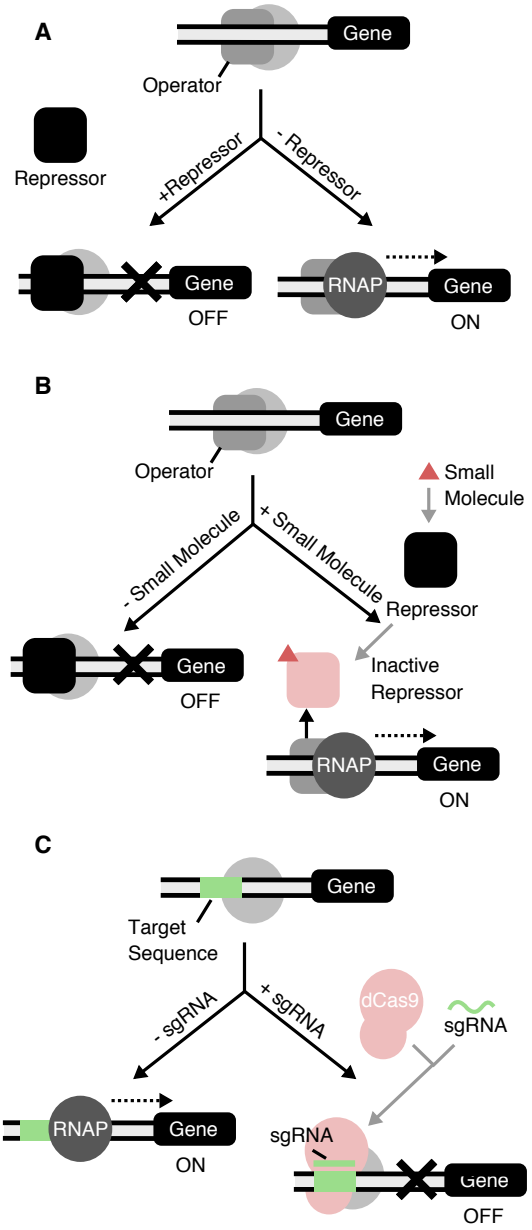


Figure 2.2: Proteins regulate transcription initiation. (**A**) Schematic shows how repressors regulate transcription. (**B**) Schematic shows how repressor based genetic sensors modify transcription in response to small molecules. (**C**) Schematic shows how the CRISPRi system can be used to regulate transcription.

makes engineering regulatory programs a challenge [124]. As the concentration of repressor protein increases, the rate of transcription initiation decreases, because fewer RNAPs can bind to the DNA (**Figure 2.2A**).

Genetic sensors are common in nature [24]. Bacteria use them to sense changes in the concentration of nutrients in their environment and to regulate the transcription of enzymes that can digest them in response [69]. A common example is the LacI repressor. LacI binds to operator sites in the P_{lac} promoter upstream of genes that encode enzymes to metabolise lactose

[37, 49, 119]. LacI is constitutively expressed, preventing these genes being transcribed until a metabolite of lactose, allolactose, is present. The conversion of lactose into allolactose is catalysed by the enzyme β-galactosidase, which is encoded in the gene LacZ, one of the genes regulated by P_{lac}. LacI interacts with allolactose causing a conformational change in the protein that reduces its affinity for the operator site DNA, allowing RNAP to more readily bind to the P_{lac} promoter and transcribe the lactose metabolising genes [38]. Thanks to this mechanism, when allolactose and therefore, lactose is present, bacteria sense it and produce enzymes to metabolise it. Using this mechanism, the required genes are only expressed when needed, conserving cellular resources. Lactose is the inducer for the promoter P_{lac}. Synthetic biologists utilize this system and many other small molecule inducible repressor-promoter pairs to allow them to regulate the transcription rates of many genes simultaneously (**Figure 2.2B**) [94, 100, 124, 128, 132]. In the case of P_{lac}, the promoter is placed upstream of the gene to be regulated. Expression of the gene of interest is controlled by adding lactose or the analogous chemical isopropyl β–d–1–thiogalactopyranoside (IPTG) to the system [132].

Meyer et al. built on this idea by taking a set of 12 naturally occurring small molecule sensors and using directed evolution to optimise their performance and reduce their interactions with each other [92]. The resulting 12 sensors all exhibit fold changes >100 and the low cross-talk between devices allows them to all be used concurrently in the same cell. They also integrated the repressor proteins into the genomes of wild type (MG1655), cloning (DH10- β) and protein expression (BL21) strains of *E. coli*. This work is significant as it enables the rate of transcription initiation for up to 12 genes to be regulated simultaneously using small molecules and the "Marionette" strains, enabling us to build more complicated genetic circuits with many inputs.

To demonstrate the utility of these strains, 5 enzymes in a pathway that produces the valuable chemical lycopene were put under the regulation of 5 sensors. An optimization strategy was employed to balance enzyme synthesis in a way that would maximize lycopene production. 972 combinations of inducer were tested, representing 972 different combinations of enzyme synthesis rates. Their approach yielded a strain that produces 90 mg/L of lycopene, a titer similar to previous best efforts [139]. If enzyme synthesis rates were balanced using traditional methods, like varying RBS and promoter strengths, 972 genetic constructs, amounting to 7 Mb of DNA, would have had to have been built. Using Marionette, optimal enzyme synthesis rates were found using a single genetic construct, an approach significantly faster and cheaper than building and testing a library of circuit variants.

A second example of where synthetic biologists have exploited native mechanisms that regulate genes are two-component systems. These consist of two proteins: a histidine kinase type enzyme and a response regulator protein (RR) [8]. The activity of the histidine kinase, is activated in response to external stimuli [77]. The active histidine kinase transfers a phosphoryl group to the response regulator, which activates it and enables it to either activate or repress a gene by interacting with a promoter sequence [66]. The way that RRs vary the rate of transcription, varies

between systems. Two component systems are useful to synthetic biologists because thousands of them have been identified in nature [62, 77]. They are therefore a large source of possible genetic sensors for synthetic biologists and their power lies in the fact they can detect a large, diverse range of stimuli like: antimicrobial peptides [65], human hormones [70] or even specific wavelengths of light [40].

Landry et al. used synthetic biology to improve a wild type sensor that made it better at detecting nitrate. They tuned the sensor's detection threshold, i.e. the concentration of inducer needed to activate their outputs, by making two different mutations to its histidine kinase – NarX. The first point mutation (C415R) targets the interface between NarX and its cognate RR – NarL , which reduces phosphatase activity and therefore the repression of transcription, increasing the sensor's sensitivity [66]. The second point mutation (D558V) targets the catalytic ATP binding domain of NarX , which more strongly reduces phosphatase activity [66]. The C415R and D558V systems were 34–fold and 381-fold more sensitive, respectively, than the threshold of the wild type system, which was 762 µM. The C415R more accurately detected nitrate over a wider range of concentrations in soil treated with commercial fertilizer [77]. It detected nitrate between concentrations of 5.62 µM and 562 µM, allowing it to detect nitrate concentrations >5-fold lower than the wild type system, which detected nitrate between 31.6 µM and 562 µM.

A third example of a synthetic transcriptional regulator that has been engineered from a natural system is the CRISPR interference system (CRISPRi) [108]. CRISPRi allows us to target and repress transcription of any gene irrespective of the target sequence. The CRISPRi system is derived from a bacterial genetic immunity pathway, characterised by sequences that feature clustered regularly interspaced short palindromic repeats (CRISPR) [93]. In the modified CRISPRi system, a single guide RNA (sgRNA) forms a complex with a catalytically inactive (or dead) form of CRISPR-associated protein 9 (Cas9), a protein that normally catalyses the formation of double stranded breaks in target DNA. The sgRNA guides the complex to a complementary target site on a DNA strand that dCas9 binds to (**Figure 2.2C**). By making the sgRNA complementary to a promoter of interest, dCas9 can be used to block RNAPs and repress transcription of downstream genes by up to 100–fold [108].

2.5.2 Tuning rate of transcription with RNA

The process of gene transcription is made up of many sub-steps. After transcription initiation, during a process called promoter escape, the RNAP holoenzyme, which is bound to the promoter sequence through its sigma factor, usually releases its sigma factor, removing its contact with the promoter and allowing the RNAP to "escape" [3, 112]. The enzyme undergoes a conformational change and the initiation complex between the RNA, DNA and RNAP is converted into a highly stable elongation complex [97]. In the next step, elongation, the RNAP-nucleic acid complex cycles between active and inactive states [3, 112]. During the active state, the RNAP opens, allowing it to bind the substrate: the correct NTP complementary to the nt on the DNA being transcribed

[3, 112]. In the RNAPs catalytic site, a phosphodiester bond forms between the NTP and the growing RNA. The by-product pyrophosphate is released and the newly bound RNA nt, along with DNA, moves 1 nt relative to the RNAP, towards the RNA exit channel [10]. This process continues until a terminator sequence is transcribed.

Intrinsic terminators halt transcription elongation independently of other molecules, whereas rho-dependent terminators depend on a separate protein to terminate transcription. Intrinsic terminators are made up of two sub sequences: a small hairpin followed by a sequence of 7 to 9 uracil (U) nucleotides [133], which form the terminator's U-tract. According to one model of intrinsic termination, the hairpin is transcribed first, but cannot form a secondary structure within RNAP [133, 136]. Once transcribed, RNA in the nascent chain remains hybridized to DNA for a period. U (RNA) and A (DNA) form the weakest base pair, so the weak hybridization between the U-tract and DNA template causes transcription to pause, allowing the hairpin to fold into its secondary structure [76]. The hairpin forms in the RNAP exit channel, causing the RNA in the U-tract and DNA template to separate, thus terminating transcription [76].

There are several genetic parts that allow us to prematurely halt transcription elongation by causing an intrinsic terminator to form in a transcript before a full coding sequence has been transcribed. For example, small non-coding RNAs (ncRNA) can interact with a specific sequence on the nascent mRNA strand. This interaction causes the RNA to form an intrinsic terminator that halts transcription before the protein coding region of the gene can be transcribed. Examples of devices that use this mechanism are RNA attenuators [84, 127] and Small Transcription Activating RNAs (STARS) [28, 29]. RNA attenuator sequences, which occur in nature, are found upstream of coding sequences. When an antisense small RNA (sRNA) is transcribed at the same time as the attenuator, the sRNA interacts with the attenuator, causing the attenuator to fold and form an intrinsic terminator (**Figure 2.3**) [11, 60]. Lucks et al. engineered a synthetic RNA attenuator based on a naturally occurring system that repressed a gene by 87% when the sRNA was highly expressed [84, 127] (**Figure 2.3**).

Chappel et al. developed a new class of regulator based on attenuators [28, 29]. In their new system, an engineered terminator sequence is placed upstream of a gene of interest, containing a sequence complementary to a STAR. When the STAR is absent the terminator forms, halting transcription. However, when the STAR is present, it binds to the terminator sequence in the mRNA, altering its secondary structure such that termination is bypassed (**Figure 2.3B**) [28, 29]. It has been shown that many orthogonal STARS can be designed using a computational approach displaying up to a 9000-fold activation in the presence of their cognate STAR [29].

2.6 Tuning the rate of translation

The information encoded in a transcript is finally converted into a protein during the process of translation. Translation is again made up of multiple steps. In bacteria, the first step is

translation initiation where the 30S ribosomal subunit along with initiation factors bind to the mRNA being translated, forming a 30S initiation complex (30SIC). Ribosomal RNA (rRNA) inside the 30SIC forms hydrogen bonds with a specific sequence of mRNA directly upstream of the start codon, contained within the RBS, known as the Shine–Dalgarno sequence (SD) [122, 141]. Next a portion of the ribosomal subunit binds to a 3 nt sequence (the start codon - AUG), signifying where translation of mRNA should start. Ribosomes then translocate 3 nt (a codon) at a time, with each of 64 possible codons mapping onto one of the 20 canonical amino acids [111, 118]. Each amino acid is attached to the growing nascent amino acid chain to produce a protein. Translation initiation is often the rate limiting step in protein synthesis. When this is the case, the rate of protein production will match the translation initiation rate [105].

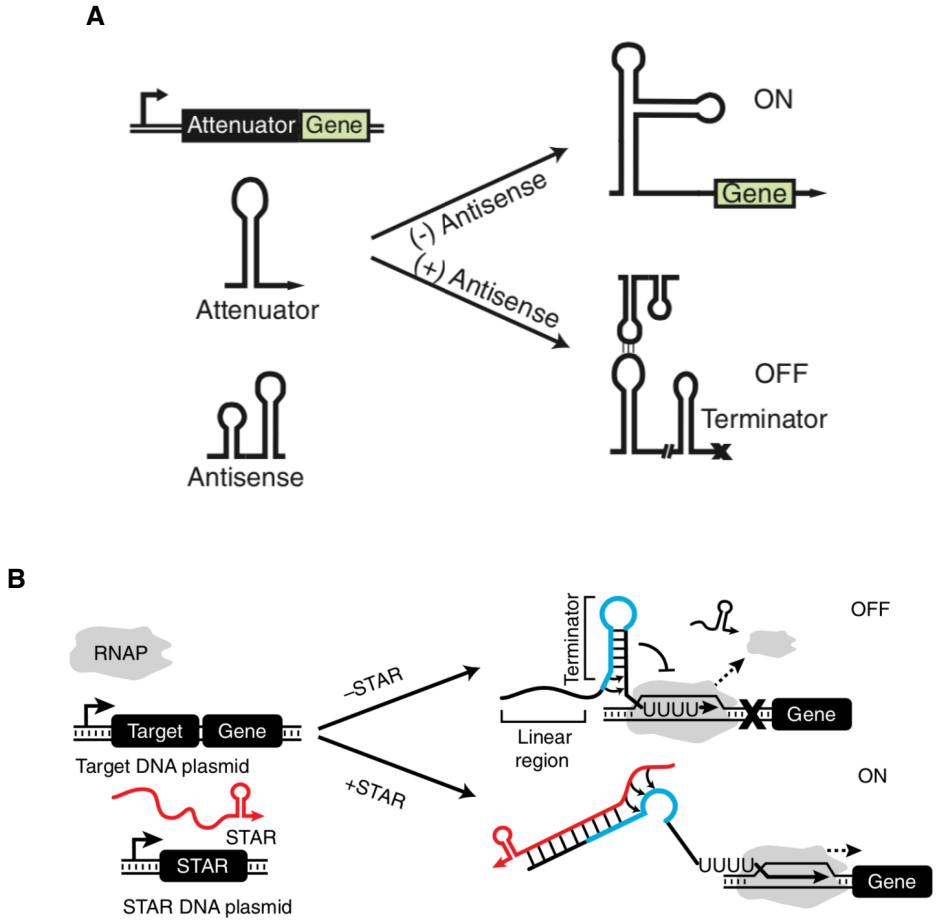


Figure 2.3: Schematics of different RNA-based transcriptional regulatory mechanisms. (**A**) The mechanism RNA attenuators use to regulate transcription (reprinted from *Nucleic Acids Research*, 41, Takahashi & Lucks, A modular strategy for engineering orthogonal chimeric RNA transcription regulators, 7577–7588, (2013) [127]) (**B**) The mechanism Small Transcription Activating RNAs (STAR) use to regulate transcription (reprinted from *Nat. Commun.*, 8, Chappell et al., Computational design of small transcription activating RNAs for versatile and dynamic gene regulation, 1051, (2017) [29].)

2.6.1 Modifying ribosome binding site strength

Translation is initiated at a rate related to the affinity that ribosomes have for the RBS. Several molecular interactions between subunits of the ribosome and the RBS play a major role in governing the rate of translation initiation. In the first step of translation initiation, the 16S ribosomal RNA (rRNA), a component of the 30S subunit, hybridizes to the RBS [31]. The distance between the 16S rRNA binding site and the start codon affects the translation initiation rate. Secondary structures at or around the RBS also affect the rate of translation initiation [115]. Therefore, the translation initiation rate can be tuned by changing the sequence of the RBS in relation to these various features.

Salis et al. built an equilibrium statistical thermodynamic model that predicts the translation initiation rate for a RBS sequence based on the strength of binding between the RBS and the 30S ribosome complex [115]. To validate how well the model predicted translation initiation rate based on RBS sequence, a library of 28 existing RBSs was used to express a red fluoresecent protein (RFP). There was a linear relationship between the predicted binding affinity (representative of relative translation initiation rates) and fluorescence. The correlation, when fit using linear regression, yielded a squared correlation coefficient (R^2) of 0.54 [115]. Building on this, to forward engineer novel RBSs, the thermodynamic model was coupled with a simulated annealing optimization algorithm that predicts RBS sequences that have a specific binding affinity for ribosomes that are closest to a user defined value. They used this approach to design 29 synthetic RBSs that they used to drive RFP expression experimentally. Predicted binding affinities and therefore relative protein production rates correlated well with fluorescence values (R^2 = 0.84).

2.6.2 Regulating the rate of translation using RNA-based parts

Naturally occurring riboregulators form secondary structures near or around RBS sequences to regulate the translation initiation rate of a gene [16, 17, 46, 53–55, 67, 68, 79, 88]. The secondary structures in riboregulators vary in strength, allowing for the rate of translation to be controlled over a wide range. In most riboregulators, a separate complementary antisense RNA can hybridize with the riboregulator and affect the strength of the secondary structure at the RBS [46, 79]. In some cases, antisense RNAs use this mechanism to reduce translation initiation. For example, antisense RNAs can bind to riboregulators to form stronger secondary structures around the RBS, causing a reduction in the rate of translation initiation. Conversely, some antisense RNAs activate translation initiation, binding to their riboregulator and causing them to lose their secondary structure. This enables ribosomes to access the RBS and increase the translation initiation rate.

Isaacs et al. engineered the first synthetic riboregulator [68]. The first component is placed upstream of the gene of interest. It is a 19 nt sequence that contains: the reverse complement of the 5 nt found upstream of the RBS, the RBS sequence, and the 6 nt downstream of the RBS. This sequence of RNA binds to the RBS and its flanking sequences, preventing ribosomes from

binding to the RBS. The transcribed mRNA cannot be translated and so is called a cis-repressed RNA (crRNA).

The second component is another RNA molecule called a trans-activating RNA (taRNA). The taRNA contains the reverse compliment of the crRNA. When the crRNA is transcribed on its own, ribosomes are blocked from the gene's RBS, so the gene of interest is not translated. When taRNA and crRNA are transcribed at the same time the taRNA binds to the crRNA and exposes the RBS, activating translation (**Figure 2.4A**). Conversely, examples exist in nature where an independent RNA molecule, or antisense RNA, represses translation of a gene of interest. In these cases, the antisense RNA binds to the translation initiation site of the gene of interest, blocking ribosomes from translating the gene [68].

An example of a riboregulator that is repressed by an antisense RNA is the RNA-IN-RNA-OUT (RNA-IN/OUT) system. An antisense RNA sequence called the RNA-OUT contains a sequence that is the reverse compliment of a regulating region of RNA upstream of the gene of interest (RNA-IN) [23, 73, 86]. The translation initiation site, including the RBS and the start codon (AUG), is contained within the RNA-IN sequence. When the gene of interest is transcribed on its own, ribosomes can bind the translation initiation site and translate the gene of interest. When RNA-OUT is transcribed at the same time, it binds the complimentary RNA-IN region blocking ribosomes and thus, turning off translation of the gene.

Mutalik et al. built a library of synthetic RNA-IN/OUT pairs [98]. They made mutations to a

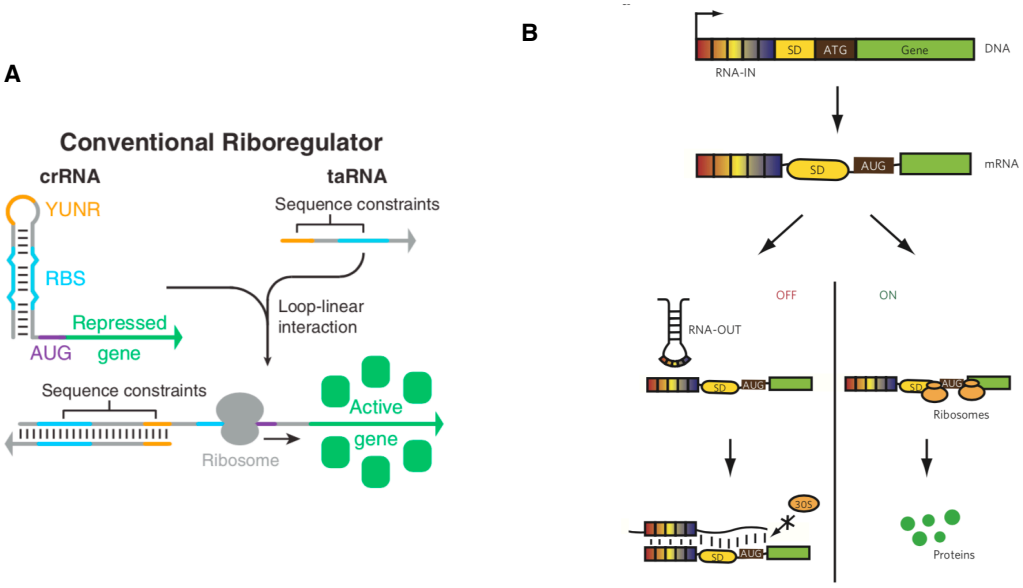


Figure 2.4: Schematic of how several RNA devices regulate translation. (**A**) Mechanism that riboregulators use to regulate translation (reprinted from *Cell*, 159, Green et al., Toehold Switches: De-Novo-Designed Regulators of Gene Expression, 925-939, Copyright (2014), with permission from Elsevier). (**B**) Mechanism RNA-IN/OUT systems use to regulate translation (reprinted from *Nat. Chem. Biol.*, 8, Mutalik et al., Rationally designed families of orthogonal RNA regulators of translation, 447–454, Copyright (2012), with permission from Springer Nature).

wild type system and tested the binding of these variants *in-vitro*. They used the results of this study to parameterise a model that allowed them to predict the behaviour of new sequences. They then built and tested a large set of these variants to validate their model [98] (**Figure 2.4B**).

Both riboregulators [68] and RNA-IN/OUT systems [98] provide synthetic biologists with parts to engineer more complicated genetic devices such as logic gates [16, 98]. Despite the clear value of these devices, there are still problems with their designs. For example, both contain sequences of RNA that bind directly to the RBS and start codon of the gene of interest. When designing new riboregulators, their sequences must include all or part of the RBS sequence and start codon. This puts a constraint on riboregulator design, limiting the number of functional sequences that can be designed [59]. A second limitation is that the performance of these RNA-based parts is normally relatively poor, with the fold-change between on and off states often much less than 100-fold (less than 55-fold for synthetic riboregulators [17] and less than 10-fold for the RNA-IN/OUT devices [98]).

Green et al. overcame these problems by engineering a new class of riboregulator called a toehold switch (THS) [54, 55]. THSs do not include sequences of RNA that bind directly to the RBS or a start codon in their design. They also give up to a 400-fold activation of gene expression. The system consists of two components: the THS, a ~100 nt long sequence that forms a secondary structure around a strong RBS to prevent translation and a sRNA, a ~65 nt long sequence that binds to a complementary portion of the THS to activate translation.

Once a toehold switch is transcribed with its downstream gene, the switch RNA folds forming a hairpin loop secondary structure that sequesters the RBS, preventing ribosomes from binding it. Unlike previous riboregulators, the RBS is not bound by any bases in the secondary structure. The switch RNA also contains a toehold, a single stranded overhanging sequence of RNA at the bottom of the hairpin loop's stem (**Figure 2.5**). When sRNA is transcribed at the same time as switch RNA, it initially binds to the toehold, triggering a branch migration process that ends when the sRNA binds to its complementary sequence in the switch's stem. sRNA binding changes the conformation of the THS's secondary structure, exposing the RBS and enabling translation. The relative concentration of trigger to switch RNA therefore determines the proportion of mRNAs that can be translated. The amount of energy needed to unfold the hairpin loop of the switch structure can be predicted using a thermodynamic model, opening up the ability to efficiently design THSs computationally [54, 55].

2.7 Environmental conditions affect circuit performance

2.7.1 Circuit performance depends on the host

A genetic circuit's performance can change depending on where it is used. The species or strain of cell that acts as a host and environmental conditions such as temperature, concentration of nutrients and oxygen all affect how biological parts in a circuit behave and thus its overall

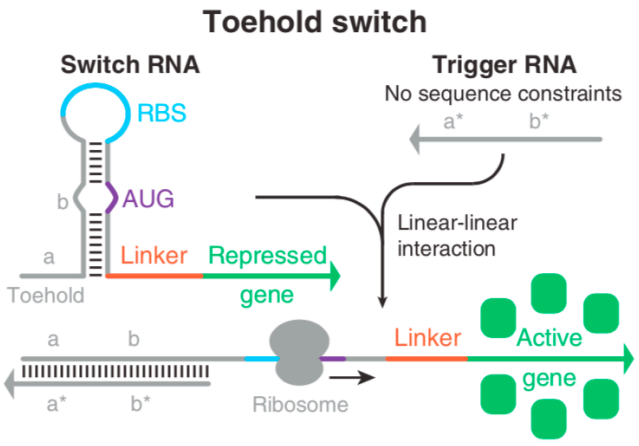


Figure 2.5: Mechanism toehold switches use to regulate translation. Switch RNA forms a secondary structure that prevents ribosomes binding the sequestered ribosome binding site (RBS) and start codon (AUG) (reprinted from *Cell*, 159, Green et al., Toehold Switches: De-Novo-Designed Regulators of Gene Expression, 925-939 Copyright (2014), with permission from Elsevier.

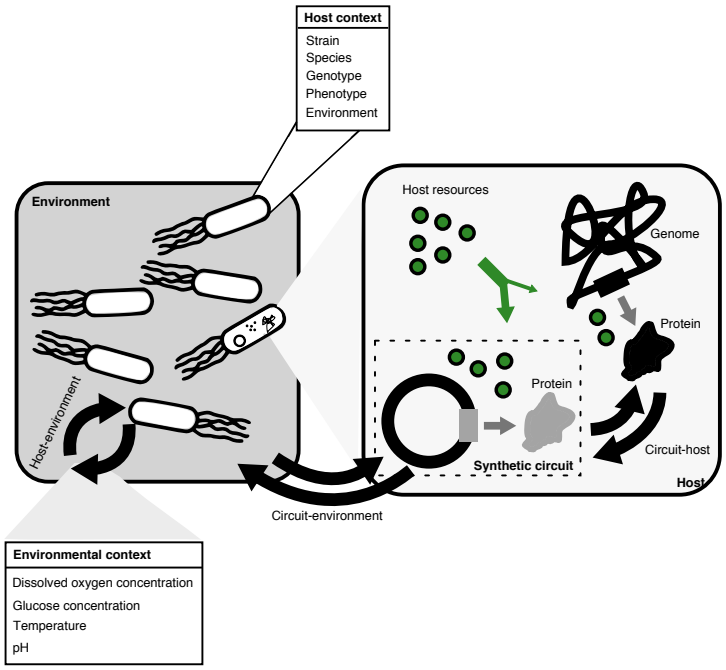


Figure 2.6: Schematic shows how environment and host interact with synthetic circuits. Expressing a synthetic circuit sequesters host resources needed for protein production, leaving fewer resources for host function (green arrow). Grey arrows represent protein production.

function [21].

Gene's that are part of a synthetic circuit rely on the host cell's resources, such as RNAPs, ribosomes and amino acids, to function. Different species and strains naturally produce different amounts of these resources and alter their availability in response to changes in the environment. As such, the rate that proteins used in circuits are produced will vary depending on the host cell

and environment the circuit is used in. These contextual effects of the host and environment can significantly impact the behaviour of genetic parts and cause circuits to behave unexpectedly and even fail [21, 52] (**Figure 2.6**).

While the host can affect the function of a genetic circuit, a genetic circuit itself can affect the host. For example, when non-native (heterologous) proteins are produced, they sequester shared protein expression resources from the cell [57, 109, 135], reducing those available to carry out essential functions, such as metabolism and cell maintenance. If this burden is significantly high, genes associated with stress are activated. These genes have evolved to help organisms survive harsh conditions, but come at the cost of often reducing gene expression globally. For example σ^{32} globally up-regulates transcription of many genes simultaneously during heat shock stress [113]. The global expression of this large set of genes sequesters resources needed to produce proteins, further impacting the performance of any circuit present [27].

2.7.2 Circuit performance depends on environmental context

In addition to the effect of the host cell, circuit performance also varies depending on the environment. This affects circuit performance in two major ways. First, it can impact host cell physiology. Second, changes in the environment can directly affect the contents of the cell and the essential chemical reactions of a circuit. For example, a small molecule might trigger a sensor due to a similar chemical structure.

Factors such as temperature, nutrients and dissolved oxygen all change the rate that cells grow at [52, 95]. Growth rate, in turn, affects the rates of cellular processes such as transcription and translation [42]. Changes in the environment also cause the host cell to adapt for survival [121]. For example, when *E. coli* cells grow in an environment with a limited amount of carbon, proteins are produced that break down molecules made of carbon already found in the cell [140].

As well as affecting circuit performance by altering the host context, environmental conditions directly affect how components in a circuit work. For example, the way that some parts perform is directly dependent on temperature. DNA and RNA parts are often required to fold into specific secondary structures which determine their function. Secondary structure is highly dependent on temperature and so minor changes can have major affects on performance [30, 33].

2.7.3 Circuits perform differently in bioreactors

Parts used in genetic circuits are generally chosen based on characterisation data that provides an input-output response function. This characterisation is often performed under homogeneous lab conditions, at a temperature of 37 °C, in a low volume of nutrient rich media and with high concentrations of dissolved oxygen [100]. This approach has its advantages. Characterising components under standard conditions makes it easier to generate reproducible data and allows components built in different labs to be easily compared. This expands the number of components available to use in new designs. However, if these parts are used in conditions that differ from

those used during characterisation, they will likely perform differently. This is problematic when genetic circuits need to be used outside of the lab, where there is the possibility of exposure to a large range of environmental conditions.

Genetic circuits can be used to engineer microbe strains that more effectively produce valuable chemicals [95, 142]. Strains used to produce chemicals in industry are grown in bioreactors with large volumes (up to 250,000 L for processes using bacteria), to improve the productivity of the processes. When a production strain is being transferred from low volume lab scale to a large scale reactor, problems often arise. Strains optimized for lab conditions, do not generally perform well at industrial scales [80].

The reason for this is that within a bioreactor, growing cells are exposed to a wide range of conditions. As the volume of a bioreactor increases, the amount of power needed to stir the reactor becomes physically limited, resulting in the liquid culture not fully mixing [36, 58, 121]. In comparison, the low volumes of liquid used to grow strains in the lab mix easily, so environmental conditions are homogeneous and well-defined. Incomplete mixing in large bioreactors leads to the formation of different compartments that do not mix with each other. This causes nutrients to become localised and not distributed homogeneously [1] (**Figure 2.7**) .

Depending on the position of a compartment in a reactor, different environmental conditions will be experienced. Temperature, concentration of substrate, concentration of dissolved oxygen and pH will all vary. The differences in environmental conditions cause microbes to grow at different rates and trigger different cellular responses as they adapt to the conditions [140].

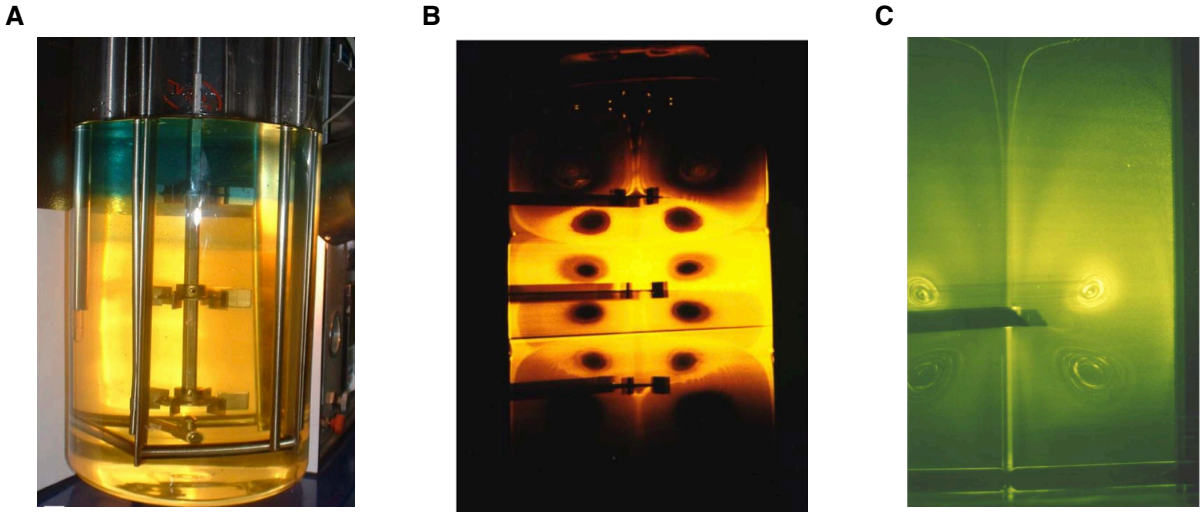


Figure 2.7: Incomplete mixing in bioreactors causes compartments with different environments to form. (**A**) Photo of a bioreactor with 3 impellers. (**B**) Compartments in the bioreactor can be seen in this 2-D laser visualisation of the bioreactor shown in panel A. (**C**) Solid particles are used to emulate cells in a stirred tank reactor with single impeller. Brighter regions are regions of high particle concentration which is analogous to high cell density (figures adapted from Alvarez et al., *Chemical Engineering Science* 60:2449-2457 2005 [1]).

Triggering these genes alters how cells metabolise nutrients and has a knock on effect on any genetic circuits they harbor [36].

Cells cultured in a bioreactor will transition between these compartments, causing their environmental conditions to vary over time. To illustrate this effect, Simen et al. exposed *E. coli* cells to the conditions experienced in two different compartments [121]. Cells were grown in a well-mixed reactor, abundant in nutrient. A fraction of the culture was continuously pumped into a second reactor where cells were starved of a nutrient. This was done for two different nutrients: glucose and ammonia. Passing cells between the two reactors emulated the passage they might experience between two compartments within a large, poorly mixed bioreactor.

To assess the response of the cells, RNA sequencing was used. They found that depriving cells of glucose or ammonia for just 110 s caused a notable stress response [121]. Genes transcribed in response to a stress response signalling molecule, 5´-diphosphate 3´-diphosphate guanosine (ppGpp), were up regulated in cells deprived of either glucose or ammonia. The ppGpp signal triggers expression of genes associated with RNA synthesis and amino acids conservation [116]. In addition to these shared effects, short term ammonia starvation up regulated genes associated with amino acid transport and biosynthesis [121] and glucose starvation down regulated genes associated with nucleotide transport and metabolism [121].

Following on from this, Simen et al. explored how fluctuations in concentration of these nutrients over a longer time scale affected cell physiology. Cultures were grown in the same two reactor system for 28 hours and RNA sequencing was performed [121]. Cells circulated between the main reactor and the secondary reactor where they were starved of either glucose or ammonia. This replicates conditions in an industrial bioreactor, where cells constantly circulate between different compartments. Cells that were intermittently starved of ammonia over 28 hours down regulated genes associated with amino acid transportation and metabolism. Cells intermittently staved of glucose up regulated genes associated with ppGpp.

To determine how fluctuations in environmental conditions affect circuit performance, we must also consider how they affect gene expression in individual cells in an entire population grown in a reactor. Delvigne et al. demonstrate this effect by showing that when cells are grown in a bioreactor, there is a large variation in the growth rate of individual cells across a population [36]. This variation in growth rate arises because the expression of genes becomes more stochastic in the dynamic environment of an industrial bioreactor [36].

The current approach of designing and characterising genetic circuits in static lab environments clearly isn't appropriate if industrial use is the goal. For genetic circuits to be of use in industry they need to be robust and adaptable to dynamic environmental conditions.

2.7.4 Designing circuits for reactors

If we model the dynamic conditions in a bioreactor, we can build better models to predict how genetic circuits will perform in real applications and improve their chance of effective scale-up.

This is not a new idea, several groups have looked into the effect that different flow regimes in a bioreactor have on the physiology of cells [36, 58, 121].

Haringa et al. modelled the movement of fluid in a bioreactor using computational fluid dynamics (CFD). They coupled the CFD model with a metabolic model of the species *Penicillium chrysogenum*, a fungus used to produce penicillin in industry [58]. This study predicted the production rate of penicillin in varying reactor volumes and configurations and showed that changing the configuration of the reactor can significantly alter the production rate of pencillin between 18% and 50%. Whilst this study demonstrated how the dynamic environment of a bioreactor can affect metabolism, the models used did not consider how changes in gene expression due to stress responses might affect production rate. At this time, no such study has been made. A novel approach for designing genetic circuits for industry would be to build on this work by using a CFD model of a bioreactor alongside a model that captures the interaction between gene expression and metabolism [102], enabling us to predict how a circuit might perform within this environment.

An alternative approach is to design adaptable circuits that are more robust to changes in host cell physiology. Ceroni et al. did this by building a genetic feedback controller that self-regulates protein production in response to the burden it places on the shared resources of the host cell [26]. First, they used RNA-seq to identify a promoter, P_{htpG1}, that becomes active when cells were burdened. Next, they built a feedback controller that regulates expression of a burdensome protein.

The controller uses CRISPRi to repress transcription of any burdensome protein by binding to its promoter. Transcription of a sgRNA that targets the burdensome promoter is regulated by P_{htpG1}, so when the cell is burdened, sgRNA is produced. The sgRNA along with dCas9, which is constitutively expressed on the same plasmid, bind to and repress transcription at the burdensome promoter. To test the controller, cells were cotransformed with the controller and a burdensome plasmid which uses an arabonise inducible promoter (P_{BAD}) to express the large the violacein biosynthesis protein (VioB), fused to a fluorescent protein (mCherry). When Viob-mCherry transcription was induced, it burdened the cell, which activated P_{htpG1} and the transcription of the sgRNA. sgRNA bound to dCas9 and the sgRNA-dCas9 complex targeted and repressed P_{BAD}, which reduced the rate of transcription of Viob-mCherry and therefore burden. After 1 hour, cells with the controller produced protein at a rate >6–fold lower than cells expressing Viob-mCherry without feedback, however, after 24 hours, cells with the the feedback controller yielded 3.9 times more protein. This is largely due to the fact that cells harbouring the feedback controller were healthier which allowed them to grow to a cell density >9–fold higher than cells without the controller, enabling more protein to be produced [26].

2.8 Summary

Humans have harnessed biology to produce valuable chemicals for thousands of years. However, technological innovations over the past century have given us more control over how we harness it. By modifying the DNA of micoorganisms, we can engineer new strains to produce chemicals more effectively. However, these developments are expensive. As such, the institutions that invest in research and development expect innovations to lead to economic growth. This means strains engineered using new technology must be more productive to ensure a return on their investment.

Innovations in biotechnology have resulted in new tools that biological engineers can use to rationally engineer microbes; an entire new discipline called synthetic biology has emerged as a result. Synthetic biologists use these enabling technologies with engineering principles to rationally engineer systems of gene regulation.

Genes regulate many of the processes that occur within cells. By rationally engineering new regulatory programs, synthetic biologists can more reliably control heterologous processes that convert feed-stocks into valuable chemicals. By precisely regulating these processes they can engineer strains that are more productive than those made using previous approaches.

Synthetic biologists often engineer gene regulation by considering gene regulatory networks as circuits that execute a desired genetic program. Like electronic circuits, genetic circuits are made up of parts and devices. Genetic circuits are designed by piecing together DNA parts and then this DNA is placed in a living cell for execution and testing. If the circuit doesn't behave as intended, another iteration of design, build, test is performed until a working variant is found.

Whilst the tools used in synthetic biology are of great value, there are still many open problems in the discipline that need addressing. In particular, reassembling circuits when they fail is expensive and time consuming. Furthermore, even if a working circuit is found, lab conditions do not generally mimic the diverse environments found in industrial processes. Thus, many circuits break when scaled-up for industry. Tuneable genetic circuits offer a solution to both these problems, allowing the function of genetic parts and devices to be dynamically modified to suit their current conditions. Such circuits open up new avenues for robust and adaptive biotechnologies and their development will be the focus of the remaining chapters.

CHAPTER 3

Materials and Methods

3.1 Introduction

In this chapter, we present the methods used to carry out the research presented in the rest of the thesis. This covers the culturing and modification of the bacteria *Escherichia coli* (*E. coli*), assembly and characterisation of genetic devices, modelling devices' behaviours and the software used in the research.

3.2 *E. coli* culturing and modification

3.2.1 *E. coli* strains

E. coli strain DH5-α (F $^-$ ϕ80$dlacZ\Delta M$15 $\Delta(lacZYA-argF)U$169 $recA$1 $endA$1 $hsdR$17(r_k^-,m_k^+) $phoA$ $supE$44 λ^- $thi-1$ $gyrA$96 $relA$1) was used for cloning. *E. coli* strain BL21 Star (DE3) (F $^-ompThsdS_B$ (r_B^-,m_B^-) gal dcm rne131 ($DE3$)) (Thermo Fisher Scientific, C601003) was used to characterize devices, unless otherwise specified.

3.2.2 Media

All water used to prepare media was ultra-purified using a water purifier (Merck-Millipore, Milli-Q Integral ultrapure water type1). Lysogeny Broth (LB) (Miller) (10 g/L tryptone, 10 g/L NaCl, 5 g/L yeast extract) (Sigma-Aldrich, L3522) and LB Broth (Miller) with agar (15 g/L agar, 10 g/L tryptone, 10 g/L NaCl, 5 g/L yeast extract) (Sigma-Aldrich, L3147) were used for cloning. Dry media was dissolved in water then autoclaved at 121 °C for 15 min to sterilise.

Supplemented M9 minimal media (M9 media salts (6.78 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 1 g/L NH_4Cl, 0.5 g/L NaCl; Sigma-Aldrich, M6030), 0.34 g/L thiamine hydrochloride (Sigma-Aldrich, T4625), 0.4% D- glucose (Sigma-Aldrich, G7528), 0.2% casamino acids (Acros, AC61204-5000),

2 mM $MgSO_4$ (Acros, 213115000), and 0.1 mM $CaCl_2$ (Sigma-Aldrich, C8106)) was used to characterise devices unless stated otherwise. M9 salts were autoclaved at 121 °C for 15 min to sterilise. All other components were sterilised by filtration using 0.22 µM sterile PES syringe filters (Star Lab, E4780-1226). Antibiotics were added to media to select for resistance markers in relevant plasmids.

3.2.3 Antibiotic stocks

To make antibiotic stocks, antibiotics solids were dissolved in ultra-pure water then sterile filtered through 0.22 µM sterile PES syringe filters (Star Lab, E4780-1226). 1000x (50 mg/mL) ampicillin trihydrate (Sigma, A1593-25G) and kanamycin sulfate (Sigma-Aldrich, K1637) were made and stored at -20 °C. Stocks were thawed and diluted to working concentration (50ng/mL) for both ampicillin and kanamycin.

3.2.4 Glycerol stocks

Glycerol stocks of bacteria strains were made by mixing overnight liquid culture with 60% (v/v) glycerol, to give a final concentration of 20% glycerol, and placed in a -80 °C freezer.

3.2.5 Conditions for growing cells

Cells grown on agar plates were grown at 37 °C for 16 hours. Cells grown in liquid media were grown in culture tubes at 37 °C and 250 RPM in an orbital shaking incubator (Stuart, SI500). When cells were grown in 96 well plates, plates were covered with a breathable seal (Star Lab, E2796-3015) and grown at 37 °C and 1250 RPM in a microtitre plate shaker incubator (Stuart, SI505) for 16 hours when grown overnight, 3 hours when cells from overnight culture were diluted and grown pre-induction or 5 hours when cells were induced.

3.2.6 Transformation of *E. coli*

To transform *E. coli* DH5-α cells we thawed a tube of competent cells on ice until defrosted (around 10 min). 50 µL of thawed cells was pipetted into a microfuge tube then 1-5 µL of plasmid DNA with a mass of 1 pg – 100 ng was pipetted into thawed cells. Cells were mixed by flicking the tube 5 times. The cell and plasmid mix was incubated on ice for 30 min. Next, cells were placed in a dry heating block (Thermo Fisher Scientific, 88870004) at 42 °C for 30 seconds then incubated on ice for 5 min. 950 µL of SOC warmed to room temperature was added to the cells which were then incubated at 37 °C with a rotation speed of 1250 RPM in a microtitre plate shaker incubator (Stuart, SI505) for 60 min. 50 µL of transformed cells was pipetted onto an agar selection plate and spread over the surface. The cells were incubated overnight at 37 °C. To transform One Shot *E. coli* BL21 Star (DE3) cells (Invitrogen, C6010-03), the same procedure was used, however, one tube of cells was thawed on ice and 1–5 µL of plasmid, at a concentration

of 5–50 ng/µL was pipetted into thawed cells. Once the cells were transformed, 20 µL and 200 µL of cells were pipetted and spread onto two separate selection plates. Plates were grown overnight.

3.3 Device construction

3.3.1 Plasmid sequences

Sequences for all plasmids built in this work can be found in Appendix A.1.

3.3.2 Genetic part sequences

Sequences for all genetic parts used in this work can be found in **Table A.1**.

3.3.3 Plasmid preparation

Cells picked from single colonies on agar plates were used to inoculate 5 mL of media. Cells in liquid culture were grown for 16 hours at 37 °C at 250 RPM in an orbital shaking incubator (Stuart, SI500). 4 mL of cells were spun down for plasmid preparation. All plasmid preparation was performed using the NEB Miniprep kit (New England Biolabs, T1010S) and the standard protocol for the kit.

3.3.4 Polymerase chain reaction

Polymerase chain reactions (PCRs) were performed in a thermal cycler (Applied Biosystems, SimpliAmp). Components were gently mixed before placing in the thermal cycler. NEB Phusion high-fidelity DNA polymerase (New England Biolabs, M0530) was used to PCR fragments for cloning, using the standard protocol. Components show in **Table 3.1** were assembled on ice and PCR was performed with the thermal cycler routine shown in **Table 3.2**. Colony PCRs were performed using Quick Load Taq polymerase (New England Biolabs, M0271L) with its standard protocol. Reaction components shown in **Table 3.3** were assembled on ice before being transferred to the thermal cycler, where the routine shown in **Table 3.4** was used.

3.3.5 Gibson assembly

Gibson assembly was used to build plasmids containing the tuneable expression system (TES) (pVB001) and the tuneable NOT gate (pVB002) with the RiboJ insulator. It was performed using NEB Gibson Assembly Master Mix (New England Biolabs, E2611S). Components listed in **Table 3.5** were assembled on ice. 50-100 ng of linearized vector with a mass of insert 2-3 times that of the vector was used. For inserts less than 200 bps, 5-times the mass of the vector was used. The total volume of PCR fragment used in the reaction was less than 20% of the total volume of the reaction. Samples were incubated in a thermal cycler at 50 °C for 15 min when 2-3 fragments

Table 3.1: Reaction components for polymerase chain reaction with Phusion polymerase

Component	25 µL Reaction	Final Concentration
Nuclease-free water	To 25 µL	
5X Phusion GC buffer	5 µL	1X
10 mM dNTPs	0.5 µL	200µM
10 µM Forward Primer	1 µL	0.5 µM
10 µM Reverse Primer	1 µL	0.5 µM
Template DNA	Variable	1 pg to 10ng
Phusion DNA Polymerase	0.25	1.0 units/50 µL PCR

Table 3.2: Thermal cycler routine for polymerase chain reaction with Phusion polymerase

Step	Temperature °C	Time
Initial denaturation	98 °C	30 seconds
25-35 Cycles	98 °C	5-10 seconds
	45-72 °C	10-30 seconds
	72 °C	15-30 seconds per kb
Final Extension	72 °C	5-10 min
Hold	4 °C	

Table 3.3: Reaction components for polymerase chain reaction with Quick-Load Taq polymerase

Component	25 µL Reaction	Final Concentration
Nuclease-free water	To 25 µL	
10 µM Forward Primer	0.5 µL	0.2 µM
10 µM Reverse Primer	0.5 µL	0.2 µM
Template DNA	Variable	< 1000 ng
Quick-Load Taq 2X Master Mix	12.5 µL	1X

were used or 60 min for 4-6 fragments. After incubation, plasmids for transformation were stored at -20 °C. Plasmids were transformed into DH5-α competent cells as outlined in section 3.2.6.

Table 3.4: Thermal cycler routine for polymerase chain reaction with Quick-Load Taq polymerase

Step	Temperature °C	Time
Initial denaturation	95 °C	30 seconds
25-35 Cycles	95 °C	5-10 seconds
	45-68 (°C)	10-60 seconds
	68 °C	1 minute per kb
Final Extension	68 °C	5 min
Hold	4 °C	

Table 3.5: Reaction components used in gibson assembly

	2-3 Fragments	4-6 Fragments
Total amount of fragments	X µL of 0.02-0.5 pmols	X µL of 0.2-1 pmols
Gibson Assembly Master Mix (2X)	10 µL	10 µL
Deionized water	10 - X µL	10 - X µL
Total Volume	20 µL	20 µL

Table 3.6: Reaction components used in golden gate assembly

	Assembly Reaction
Backbone, 75 ng/µL	1 µL
Inserts	2:1 Molar ratio
T4 DNA Ligase Buffer (10X)	2 µL
NEB Golden Gate Assembly Mix	1-2 µL
Nuclease-free water	to 20 µL

3.3.6 Golden gate assembly

Golden gate assembly was used to remove the RiboJ insulator from the TES and NOT gate plasmids, yielding plasmids pVB003 and pVB004, respectively. Golden gate assembly was performed using the NEB Golden Gate Assembly Kit (BsaI-HF®v2) (New England Biolabs, E1601). Components listed in **Table 3.6** were assembled on ice. Components were incubated in a thermal cycler at 37 °C for 5 min, then heated to 60 °C and incubated for 5 min, to allow for the digestion of any remaining BsaI sites. Assembled plasmids were transformed into DH5-α competent cells as outlined in section 3.2.6.

3.3.7 Gel electrophoresis

Gel electrophoresis was used to quality control products from PCRs and assembly methods and verify that they yielded DNA products with expected lengths. To prepare a 1% agarose gel for use in a 20 mL chamber, a screw top bottle was loaded with 0.2 g of agarose (Lonza, 98200-100) and 20 mL of TAE buffer diluted with SYBR Safe (Thermo Fisher Scientific, S33102) to 1X concentration from a 10,000X concentration stock. The lid was screwed on and the bottle was shaken to dissolve agarose. The lid was undone half a turn before microwaving the agarose at 800 W for 30 seconds. The bottle was swirled to ensure agarose was well mixed before microwaving at full power for another 30 seconds. The contents of the bottle were swirled once more then the liquid agarose was poured into a 20 mL gel casting tray with appropriate well comb inserted and left to set for 40 min at room temperature.

DNA samples were prepared by mixing 5 µL of DNA in elution buffer, containing 2–20 ng/L of DNA, with 1 µL of 6 X purple no SDS gel loading dye (New England Biolabs, B7025S). The gel in its casting tray was moved into a gel box (Bio-Rad, Mini-Sub Cell GT). The well comb was removed then the gel box was filled with TAE buffer until buffer height was slightly higher than the surface of the gel. Using a pipette, DNA was loaded into relevant wells. The gel box lid was put on and connected to a power pack (Bio-Rad, PowerPac Basic). Gels were run at 300 V with a constant 80 mA current for 30 min. Gels were visualised and photographed on a gel doc (UVP, BioDoc-It).

3.3.8 Tuneable expression system assembly

The tuneable expression system was assembled into the plasmid pVB001 (**Figure 3.3**). It was built from 3 linear DNA fragments assembled using Gibson assembly (**Figure 3.1**). The first fragment, an expression cassette for small RNA (sRNA) 20 [55] was amplified from a synthesized DNA block – VB_TS_ASS_B (**Table 3.7**) (Thermo Fisher Scientific, GeneArt), using primers PVB_1_TES_fwd and PVB_2_TES_rev to give a 183 bp sequence containing P_{tac}, sRNA-20 and terminator L3S3P11. The second fragment was amplified from a second synthesized DNA block – VB_TS_ASS_A (**Table 3.7**) (Thermo Fisher Scientific, GeneArt) using primers PVB_3_TES_fwd and PVB_4_TES_rev. The resulting fragment was a 241 bp sequence containing P_{tet}, insulator RiboJ and toehold switch-20 [55]. The third fragment was amplified from circuit plasmid pAN1720 [100] with primers PVB_5_TES_rev and PVB_6_TES_FWD to make a 5760 bp linearized backbone containing *kanR*, p15A ORI, *tetR*, *lacI*, *araC* and *yfp*. Primer sequences with annealing temperatures and extension times are shown in **Table 3.8**. Primers introduced a 17 to 21 bp overhang sequence homologous to sequences at the end of fragments that would be adjacent in the final construct, allowing the plasmid to be assembled via Gibson assembly. After assembly, DH5-cells were transformed with reaction product and grown overnight on a selection plate. Six colonies were tested using colony PCR with Quick Load Taq polymerase and primers PVB_1_TES_fwd and PVB_2 _TES_rev, using the protocol outlined in section 3.3.4.

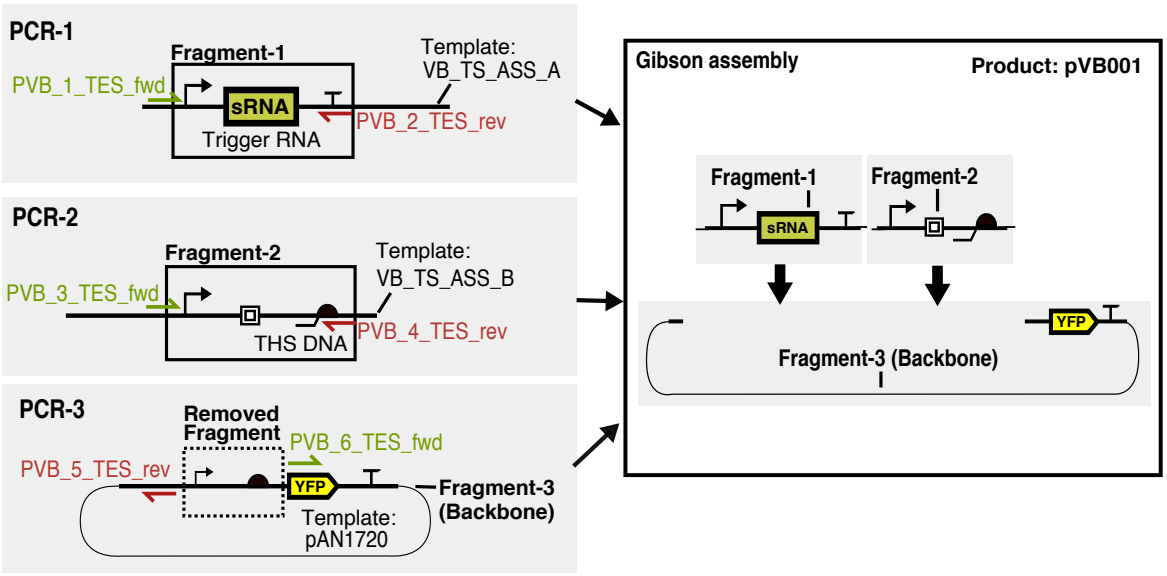


Figure 3.1: Tuneable expression system assembly Green and red half arrows show where forward and reverse primers bind to template DNA. All primers contain overhang sequences. During PCR-1 fragment–1, containing P_{tac} small RNA (sRNA) and terminator L3S3P11 was amplified from synthesized DNA block VB_TS_ASS_A. During PCR-2, fragment–2, containing P_{tet}, RiboJ insulator and toehold switch (THS) DNA fragment were amplified out of the synthesized DNA fragment VB_TS_ASS_B. During PCR-3, fragment-3 (the plasmid backbone) was amplified from plasmid pAN1720, removing the P_{BAD} promoter. Fragments are shown here in the order they were inserted into the backbone during Gibson Assembly.

Colonies containing plasmids that yielded fragments of the correct length were grown overnight in liquid media before being prepared (as outlined in section 3.3.3), and sent for sequencing using sequencing primers pSeq1_TES and pSeq2_TES (**Table 3.8**).

3.3.9 Tunable NOT gate construction

The tuneable genetic NOT gate plasmid, pVB002 (**Figure 3.3**) was assembled from 4 linear fragments using Gibson assembly (**Figure 3.2**). To produce fragment-1, pAN3938 [100] was used as a template for primers PVB_3_TES_fwd and PVB_4_TES_rev. The PCR yielded a 690 bp sequence containing PhlF and Terminator ECK120033737. To produce fragment-2, pAN4036 [100] was used as template with primers PVB_P5_NOT_fwd and PVB_P6_NOT_rev to amplify a 98 bp fragment containing a spacer and the P_{phlF} promoter. To produce fragment-3, pAN4036 [100] was used as a template with primers PVB_P7_fwd and PVB_P8_rev to amplify a 917 bp sequence containing insulator RiboJ, RBS B0064, *yfp* and Terminator L3S2P2. To produce fragment-4, *yfp* and its terminator, L3S2P21 were removed from pVB001 by PCR, using primers PVB_P1_NOT_fwd and PVB_2_TES_rev, yielding a linearized 5379 bp fragment containing: P_{tac}, sRNA-20, terminator L3S3P11, P_{tet}, insulator RiboJ and toehold switch-20. Primers and PCR conditions are shown in **Table 3.8**. Primers were designed to flank fragments with 13 to 21 bp overhang sequences. Overhang sequences were homologous to overhangs on fragments that

would be adjacent in the final construct which allowed the plasmid to be assembled via Gibson assembly (Section 3.3.5).

Assembled plasmids were transformed into DH5-α cells. 12 colonies were grown overnight and screened using the standard device characterisation protocol (section 3.4.5) in four combinations of aTc and IPTG (aTc: 0 ng/mL , IPTG: 0 µM; aTc: 0 ng/mL , IPTG: 1000 µM, aTc: 50 ng/mL, IPTG: 0 µM; aTc: 50 ng/mL, IPTG: 1000 µM). Colonies containing plasmids that showed NOT gate behaviour, i.e. had a considerably higher output fluorescence in an OFF state (aTc: 0 ng/mL , IPTG: 0 M) than an ON state (aTc: 50 ng/mL, IPTG: 1000 M) were prepared then sequenced with primers pSeq1_TES, pSeq2_TES and pSeq3_NOT which are shown in **Table 3.8**.

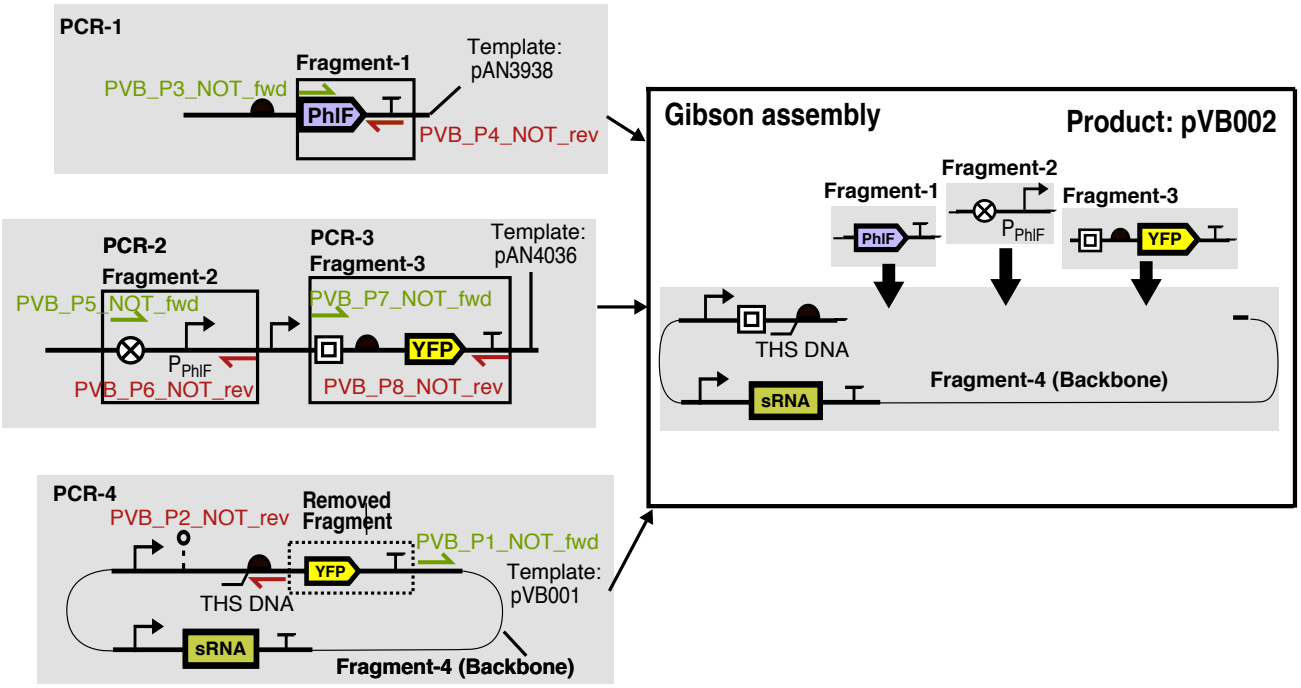


Figure 3.2: Tuneable NOT gate system assembly Schematic shows 4 PCR reactions used to amplify the fragments used to assemble the tuneable NOT gate plasmid (pVB002). Green and red half arrows show where forward and reverse primers bind to template DNA. Dotted box shows DNA removed from fragments and solid lined boxes show DNA amplified out of fragemetns. (PCR-1) Fragment-1, the PhlF open reading frame and terminator were amplified out of plasmid pAN3938 [100]. (PCR-2) Fragment-2, a spacer followed by P_{phlF} was amplified out of plasmid pAN4026 [100]. (PCR-3) Fragment-3, a YFP cassette was amplified out of plasmid pAN4026 [100]. (PCR-4), Fragment-4 (Backbone), a YFP cassette was removed from plasmid pVB001, yielding the linearized backbone used for assembly. Overhangs flanking fragments were homologous to adjacent fragments in the final construct allowed the plasmid, pVB002 to be assembled using Gibson assembly. Fragments are shown here in the order they were inserted into the backbone.

3.3.10 sRNA booster plasmid construction

The booster plasmid, pVB005 (**Figure 3.3**), used to increase the rate of transcription of sRNA was fully synthesized (Thermo Fisher Scientific, GeneArt). It contained an sRNA expression

Table 3.7: Synthesized gene fragments used to assemble tuneable expression system

Fragment	Sequence
VB_TS_ASS_A	CGCAAACCGCCTCTCCCCGCAACGATCGTTGG CTGTGTTGACAATTAATCATCGGCTCGTATAAT GTGTGGAATTGTGAGCGCTCACAATTGGGACC GTGGACCGCATGAGGTCCACGGTAAACATAAC TATAACAAGCCTACAATTCATTCAAACCCAATT ATTGAACACCCTTCGGGGTGTTTTTTTGTTTCT GGTCTACCTACTCCACCGTTGGCTTTTTTCCCT ATCAGTGATAGAGATTGACATCCCTATCAGTGA TAGAGATAATGAGCACAGCTGTCACCGGATGTG CTTTCCGGTCTGATGAGTCCGTGAGGACGAAAC AGCCTCTACAAATAATTTTGTTTAAGGGCGTTAA TCTCTGGCTTGCTTTATGTCTGTAAACAGAGGAG ATACAGAATGAAAGCAAGCAACCTGGCGGCAGC GCAAAAGATGCGTAAACGCAGCCTGAATGGCGAA TG
VB_TS_ASS_B	CGCAAACCGCCTCTCCCCGCAACGATCGTTGGCT GTGTTGACAATTAATCATCGGCTCGTATAATGTGT GGAATTGTGAGCGCTCACAATTGGGTCATGACTG GGACACGCCAGTCATGAGAATACAGACATAAAGCA AGCCAGAGATTAACGAAGCCAATTATTGAACACCC TTCGGGGTGTTTTTTTGTTTCTGGTCTACCTACTCC ACCGTTGGCTTTTTTCCCTATCAGTGATAGAGATTG ACATCCCTATCAGTGATAGAGATAATGAGCACAGCT GTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGT GAGGACGAAACAGCCTCTACAAATAATTTTGTTTAA GGGTGAATGAATTGTAGGCTTGTTATAGTTATGAAC AGAGGAGACATAACATGAACAAGCCTAACCTGGCG GCAGCGCAAAAGATGCGTAAACGCAGCCTGAATGG CGAATG

All sequences are written from 5' to 3'.

cassette, a 133 bp sequence made up of P_{T7}, sRNA-20 and a T7 terminator. The backbone of the plasmid contained an *ampR* ampicillin resistance cassette and a pColE origin of replication.

3.3.11 Removal of RiboJ insulators

RiboJ insulators were removed from the TES (pVB001) and NOT gate (pVB002) plasmids by amplifying the plasmids using forward and reverse primers upstream and downstream of the RiboJ sequence, respectively so that the linear backbone product did not contain the RiboJ sequence. Primers used in the PCRs (primers pF-RemRiboJ-THS20 and pR-RemRiboJ-THS20, **Table 3.8**) introduced BsaI sites and 4bp overhang sequences to either end of the linearized fragments. Overhangs were complimentary to overhangs on the opposite ends of the plasmid. This allowed us to perform a one pot Golden Gate reaction. Plasmids were sequenced using primers

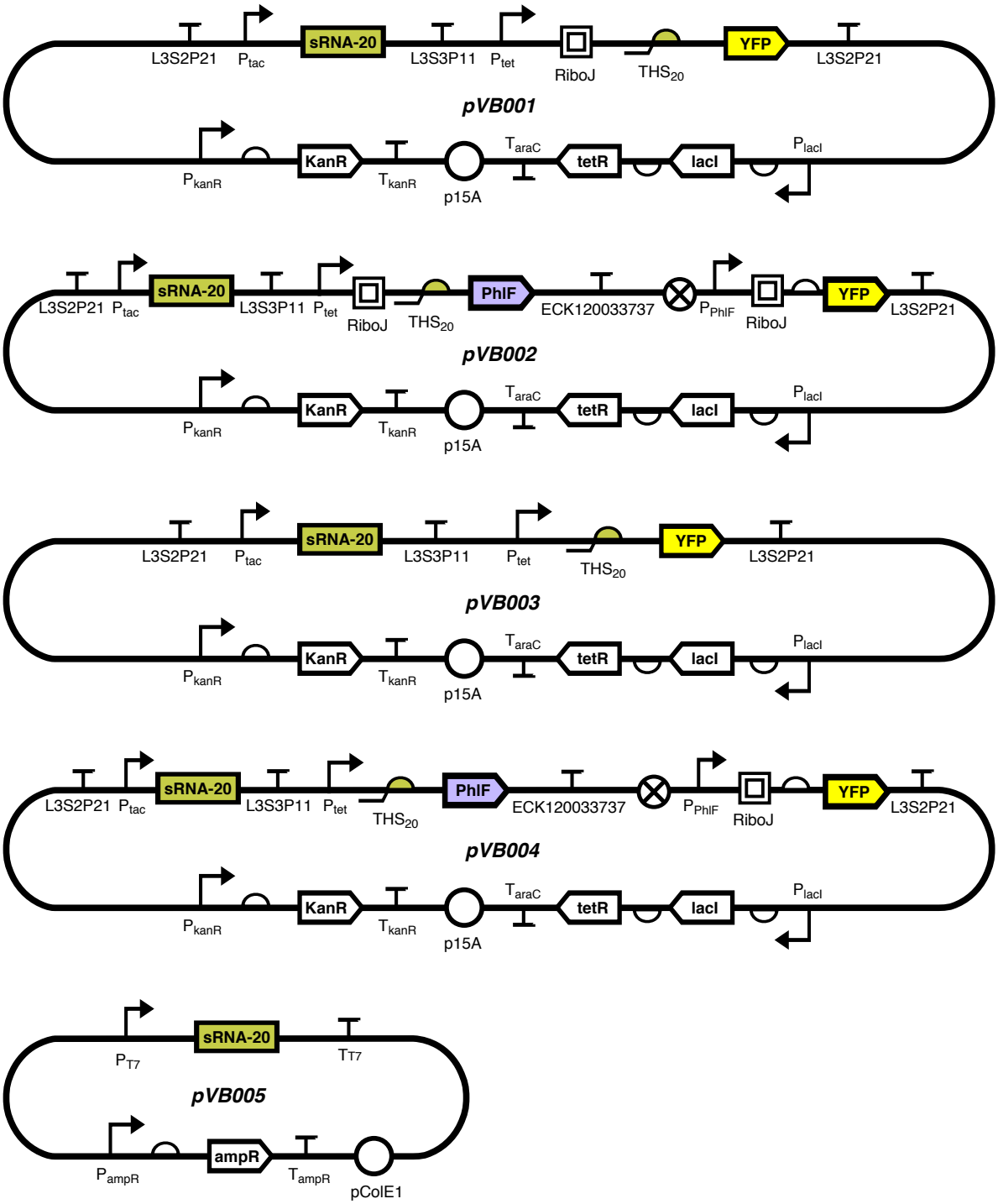


Figure 3.3: Plasmid maps. (**A**) pVB001: tuneable expression system. (**B**) pVB002: tuneable NOT gate. (**C**) pVB003: tuneable expression system without RiboJ insulator. (**D**) pVB004: tuneable NOT gate without RiboJ insulator. (**E**) pVB005: tuner sRNA booster.

pSeq1_TES, pSeq2_TES for the TES (pVB003) and pSeq1_TES, pSeq2_TES and pSeq3_NOT for the NOT gate (pVB004) (see **Table 3.8**).

3.3.12 Sequence verification

Plasmids were sequence verified using sanger sequencing (Eurofins Genomics). Sequencing primers used are shown in **Table 3.8**.

3.4 Characterising devices

3.4.1 Inducing circuits for characterisation

Single colonies from streaked plates were used to inoculate 200 µL of supplemented M9 minimal media with relevant antibiotic in a 96 well plates then grown overnight at 37 °C with a rotation speed of 1250 RPM in a microtitre plate shaker incubator (Stuart, SI505). 15 µL of the overnight culture was loaded into 185 µL of supplement M9 media. 15 µL of these diluted cells were then used to inoculate 185 µL of supplement M9 media with relevant antibiotic. Cells were grown for 3 hours, after which they were diluted 3-fold in M9 media with relevant antibiotic. 10 µL of diluted cells was used to inoculate wells loaded with 140 µL of M9 media with relevant antibiotic and inducer. For all characterizations, wells were loaded with different combinations of IPTG and aTc. For the TES, wells were loaded with combinations of IPTG (at final concentrations: 0, 10, 20, 45, 70, 100 and 1000 µM) and aTc (at final concentrations: 1, 2, 4, 8 and 50 ng/mL). For the NOT gate wells were loaded with combinations of IPTG (at final concentrations 0, 10 , 20, 45, 70, 100 and 1000 µM) and aTc (at final concentrations 0.01, 0.04, 0.14, 0.5 and 2 ng/mL). Cells were grown in 96 well plates for 5 hours before being prepared for flow cytometry.

3.4.2 Flow cytometry

Phosphate buffer saline (PBS) solution was made by dissolving a PBS tablet (Gibco, 18912014) in 500 mL of ultra-purified water which we autoclaved at 121 °C for 15 min to sterilise. In a 50 mL falcon tube 12.3 mL of PBS was mixed with 600 µL of kanamycin (50 mg/mL stock), to give a final kanamycin concentration of 2mg/mL. 90 µL of PBS mixed with kanamycin was loaded in to every well of a 96 well plate. Cells in the 96-well plate they were grown in were mixed by pipetting then loaded into the 96-well plate containing PBS and kanamycin. These were then mixed with the PBS by pipetting to fix them. YFP fluorescence of individual fixed cells was measured using an Acea Biosciences NovoCyte 3000 flow cytometer equipped with a NovoSampler to allow for automated collection from 96-well microtiter plates. Cells were excited using a 488 nm laser and measurements taken using a 530 nm detector. A flow rate of 40 µL/min was used to collect at least 100,000 cells for all measured conditions. Automated gating of events using the forward (FSC-A) and side scatter (SSC-A) channels was performed for all data using the FlowCal Python package

Table 3.8: List of primers used for plasmid assemblies, sequencing and colony PCR

Primer	Description	Sequence	Annealing temperature (°C)	Extension time (s)
PVB _1 _TES _fwd	TES assembly	ctaacgggggggccttttttttgAACGATCG TTGGCTGTGTTG	66	15
PVB _2 _TES _rev	TES assembly	aaagccaacggtggagtaGGTAGACCAG AAACAAAAAAAACACCC	66	15
PVB _3 _TES _fwd	TES assembly	gtgtttttttgtttctggtctaccTACTCCACC GTTGGCTTTTTTCC	72	15
PVB _4 _TES _rev	TES assembly	tcctcgcccttgctcacTTTACGCATCTTT TGCGCTGC	72	15
PVB _5 _TES _rev	TES assembly	caacacagccaacgatcgttCAAAAAAAG GCCCCCCGTTAG	66	30
PVB _6 _TES _fwd	TES assembly	cagcgcaaaagatgcgtaaaGTGAGCAA GGGCGAGGA	66	30
PVB _P1 _NOT _fwd	NOT assembly	ttcgttttggtccAGTTTACGGCTAGCT CAGTCCTAG	63	81
PVB _P2 _NOT _rev	NOT assembly	gctacggctcggggtacgtgcTTTACGCAT CTTTTGCGC	63	81
PVB _P3 _NOT _fwd	NOT assembly	tgcgtaaagcacgtaccccgaGCCGTAGC AGCATTGGTAG	63	15
PVB _P4 _NOT _rev	NOT assembly	ttccaccgtacgtcgaaCGTTGACACCT TTGGTCG	63	15
PVB _P5 _NOT _fwd	NOT assembly	caaaggtgtcaacgTTCGACGTACGG TGGAATC	63	15
PVB _P6 _NOT _rev	NOT assembly	atccggtgacagcttGACCTTAACGAT ACGGTACGTTTC	63	15
PVB _P7 _NOT _fwd	NOT assembly	taccgtatcgttaaggtcAAGCTGTCAC CGGATGTG	63	15
PVB _P8 _NOT _rev	NOT assembly	tgagctagccgtaaactGGACCAAAAC GAAAAAGGC	63	15
pF-RemRiboJ-THS20	RiboJ Removal	cgaa**ggtctc**aGGGCGTTAATCT CTG-GCTTGCTTTATG	60	90
pR-RemRiboJ-THS20	RiboJ Removal	ggtc**ggtctc**agcccGTGCTCATTATC TCTATCACTGATAGGGATGTC	60	90
pSeq1 _TES	TES and NOT gate sequencing	ACACAACCAATTATTGAAGGCCT CC		
pSeq2 _TES	TES and NOT gate sequencing	CAGGAGCGCACCATCTTCTTC		
pSeq3 _NOT	NOT sequencing	CGTTGTGTTATTGCAGAAG		

Underlined sequences are overlapping sequences used for assembly, lower case bases are bases introduced by the primer through PCR. Bold text is the BsaI recognition site needed for Golden Gate Assembly. All sequences are written from 5' to 3'.

version 1.2 [25] and the density2d function with parameters: channels = ['FSC-A', 'SSC-A'], bins = 1024, gate_fraction = 0.5, xscale = 'logicle', yscale = 'logicle', and sigma = 10.0.

3.4.3 Plate reader assay

Cells were grown overnight in LB media (for 16 hours) at 250 rpm in culture tubes. 10 µL of overnight culture was used to innoculate 140 µL LB media with the desired concentration of inducer into a black 96 well plate with optically clear flat bottomed wells (Brooks Life Sciences, 4ti-0223). Wells were loaded with combinations of IPTG (at final concentrations: 0, 18, 90, 200, 300 ,450, 670 and 1000 µM) and aTc (at final concentrations: 0, 0.04, 0.25, 0.63, 1.0, 1.6, 2.1, 4.0 ng/mL). OD600 and fluorescence (at excitation wavelength 512 nm and emission wavelength 529 nm) were measured from the top and bottom of each well every 16 min, respectively.

3.4.4 Autofluorescence normalisation

Plasmid pAN1201 [100] was used to measure autofluorescence. pAN1201 had the same backbone as we used to test the devices and like the device plasmids, it contained a constitutively expressed KanR, and TetR and LacI regulated by a P_{lacI} promoter but no YFP casette. We transformed pAN1201 into *E. coli* BL21 star (DE3) and grew cells in the same conditions as our devices. The fluorescence of the autofluorescence strain was measured using flow cytometry (section 3.4.2). The fluorescence we detected in cells transformed with pAN1201 was the autofluorescence of the cells. We calculated the median fluorescence of three gated populations of cells containing pAN1201 and took the mean of these values. This value was then subtracted from the median fluorescence of our devices to give an autofluorescence normalised fluorescence.

3.4.5 Characterising sensor outputs in relative promoter units

To relate the sensor inputs (concentrations of inducer), to their outputs (promoter activities), which are our device inputs , the outputs from sensors P_{tet} and P_{tac} were characterised using sensor characterisation plasmids pAN1718 and pAN1719, respectively [100]. These plasmids used the same backbone as our devices and contained their respective sensor upstream of the same YFP cassette (RiboJ, RBS–B0064, *yfp*, terminator–L3S2P21). Output fluorescence from a third plasmid (pAN1717) that used the same YFP casette regulated by a constitutive promoter (J23101), was used as reference to convert the output from the sensors (fluorescence data in arbitrary units – a.u) into relative promoter units (RPU). All plasmids were transformed into cells and grown in the same way as our devices. Cells containing pAN1718 (P_{tet}) were grown in aTc concentrations: 0, 0.01, 0.04, 0.14, 0.5, 1, 2, 4, 8, 50 ng/µL. Cells containing pAN1719 (P_{tac}) were grown in IPTG concentrations: 0, 0.5, 5, 50, 100, 200, 450, 1000, 4000, 16000 mM. Cells containing pAN1717 were grown in no inducer. For each set of strain and condition, three biological replicates were performed. Fluorescence was measured using flow cytometry and the

median fluorescence of gated populations for replicates was averaged. We standardised the output promoter activities by converting them from fluorescence into relative promoter units using:

$$RPU = \frac{\langle YFP \rangle - \langle YFP \rangle_0}{\langle YFP \rangle_{RPU} - \langle YFP \rangle_0} \tag{3.1}$$

where $\langle YFP \rangle$ is the median fluorescence of the device being measured, $\langle YFP \rangle_{RPU}$ is the median fluorescence of cells transformed with the RPU plasmid and $\langle YFP \rangle_0$ is the median fluorescence of cells transformed with an autofluorescence control. We plotted promoter outputs in RPUs against inducer concentrations to make a calibration curve (**Figure 4.6**), which was used to convert sensor input concentrations into promoter activities in RPUs.

3.4.6 Calculating intersection

For two overlapping population distributions, the fraction of individual cells from each distribution that overlaps, known as the intersection, was calculated using:

$$H(x,y) = \sum_{i=1}^{n} \frac{min(x_i, y_i)}{x_i}, \tag{3.2}$$

where $H(x,y)$ is the intersection between distributions x and y. Distributions x and y are divided into n bins that correspond to identical ranges of values, with x_i and y_i denoting the value of bin i for distributions x or y, respectively [126].

3.4.7 Computational analyses and response function fitting

Data analysis was performed using custom software written in Python version 3.6.6. Response functions for the tuneable expression system were fitted to the hill function:

$$y = y_{min} + (y_{max} - y_{min}) \frac{x^n}{K^n + x^n}, \tag{3.3}$$

where y is output YFP fluorescence (in arbitrary units), y_{min} and y_{max} are the minimum and maximum YFP fluorescences (in arbitrary units), x is the input promoter activity (in RPU), K is the input promoter activity at which the output is at half of its dynamic range (i.e.$y_{min} + 0.5 \times (y_{max} - y_{min})$, in RPU) and n is the hill coefficient. Response functions for the NOT gate were fitted to the hill function:

$$y = y_{min} + (y_{max} - y_{min}) \frac{K^n}{K^n + x^n}. \tag{3.4}$$

All curves were fit using non-linear least squares and SciPy package version 1.3.0 with the optimize.curve_fit module.

3.4.8 Finding growth rates

Growth curves were fit using an empirical model of the form:

$$y(t) = \frac{at^n}{b + t^n} \tag{3.5}$$

where t is time, y is cell density, a, b and n are empirical growth coefficient parameters [35]. The maximum growth rate was found using numerical differentiation on the curve fit using the numpy gradient function in Python version 3.6.6. The maximum gradient was the value taken as the maximum growth rate.

3.5 Modelling

3.5.1 Numerical simulations

Deterministic ordinary differential equation models were simulated using COPASI version 4.2.4 using a parameter scan that implemented a 'steady-state' task with default parameters [63] . Stochastic simulation of the biochemical models were carried out in COPASI using the tau-leap method, starting in steady state with settings: duration = 100 min, interval size = 1 min, number of intervals = 100 [63].

3.5.2 Fitting model with particle swarm optimisation

Parameterization was performed using custom software built in Python. Initial estimation was performed using parameters shown in **Table 4.1**. The scipy integrate package was used to solve the model with the integrator set to "vode", which uses an implicit Adams method to solve non-stiff problems and a backward differentiation formula method to solve stiff problems. For each input combination, the simulated trajectory for the dynamics of protein production was compared to the experimental data using least squares regression, which acted as the objective function for the optimization. A particle swarm optimization algorithm implemented in python with the pyswarm package with parameters: maxiter=100, args=dataSet, omega=0.5, phip=0.7, phig=0.4, swarmsize = 200, was used to minimize the objective function.

3.5.3 Predicting RNA binding and secondary structure

NUPACK version 2.2 v was used to simulate RNA binding and secondary structure. We simulated folding of a toehold switch (switch) and cleaved RiboJ bound to switch (cRiboJ-Switch), with parameters: nucleic acid=RNA, temperature=37°C, strand species = 1 and a maximum complex size = 1. The concentration of toehold switch mRNA was set to 5×10^4µM.

Binding and folding of the toehold switch with trigger sRNA (trigger) and cRiboJ-switch with trigger sRNA were simulated using parameters: nucleic acid = *RNA*, temperature = 37°C, strand

Table 3.9: Sequences used to model RNA folding

Fragment	Sequence
Switch	GGGCGUUAAUCUCUGGCUUGCUUUAU GUCUGUAAACAGAGGAGAUACAGAAUG AAAGCAAGCAACCUGGCGGCAGCGCAA AAGAUGCGUAAA
Trigger	GGGUCAUGACUGGGACACGCCAGU CAUGAGAAUACAGACAUAAAGCAAG CCAGAGAUUAACGAAG
cRiboJ-Switch	UCACCGGAUGUGCUUUCCGGUCUG AUGAGUCCGUGAGGACGAAACAGC CUCUACAAAUAAUUUUGUUUAAGGG CGUUAAUCUCUGGCUUGCUUUAUG UCUGUAAACAGAGGAGAUACAGAAU GAAAGCAAGCAACCUGGCGGCAGC GCAAAAGAUGCGUAAA
RiboJ-Switch	AGCUGUCACCGGAUGUGCUUUCCGG UCUGAUGAGUCCGUGAGGACGAAACA GCCUCUACAAAUAAUUUUGUUUAAGG GCGUUAAUCUCUGGCUUGCUUUAUGU CUGUAAACAGAGGAGAUACAGAAUGAA AGCAAGCAACCUGGCGGCAGCGCAAAA GAUGCGUAAA

All sequences are written from 5' to 3'.

species = 2 and a maximum complex size = 2. The concentration of toehold switch mRNA and tuner sRNA were set to: 5×10^4 µM and 7×10^5 µM, respectively. Sequences used in simulations can be found in **Table 3.9**.

3.5.4 Predicting RNA folding speeds

Speed of RNA folding was predicted using Kinfold (Vienna-RNA-package version 1.4 [45]) with the following parameters – energy model: dangle=2, Temp= $37.0°C$, logML=*logarithmic*, Par=VRNA–1.4; move set: noShift=off, noLP=*off*; simulation: num=1, time=500.00, seed=*clock*, fpt=*on* mc=*Kawasaki*, phi=1, pbounds=0.1 0.1 2; Output: log=*kinout*, silent=*off*, lmin=*off*, cut=*20.00* [45]. We simulated folding speeds for a toehold switch, a toehold switch with an upstream RiboJ (RiboJ-Switch) and a toehold switch with an upstream cleaved RiboJ (cRiboJ-Switch). Sequences used for simulations can be found in **Table 3.9**.

3.6 Data analysis software

Custom software to analyse experimental data and generate plots was written in Python version 3.6.6. Matplotlib version 3.1.1 was used to generate plots. Scipy version 1.3.0 was used with NumPy version 1.17.0 and Pandas version 0.25.0 to process experimental data and to perform numerical routines.

CHAPTER 4

A tuneable expression system

4.1 Introduction

Synthetic circuits have been shown to improve the production of valuable chemicals [48, 96]. Industrial production of valuable chemicals happens in bioreactors, where cells are grown in dynamic environments in which conditions such as temperature and concentration of nutrient change across the volume of the reactor as well as over time. These variations in conditions affect how parts and devices behave [52, 95]. Therefore, changing environmental conditions can change the shape of response functions, causing devices to become mismatched and circuits to fail. Broken devices and circuits must be rebuilt and tested, which is costly and time consuming.

Here, to address this problem we designed a tuneable expression system (TES) whose response function can be dynamically tuned by simultaneously controlling the rates of transcription and translation of a gene of interest. Using a TES removes the need to rebuild circuits when they fail. Instead, the TES can be used to tune the behaviour of individual parts until a working combination is found. This approach allows us to bypass the expensive and time consuming task of assembling new genetic designs from scratch for each DBTL iteration.

In Section 4.2 we define the tuneable expression system (TES) and derive a deterministic ordinary differential equation (ODE) model that we use to predict the device's behaviour. Then, in Section 4.3 we design a TES from real genetic parts. We adapt the parameters in our model to make them realistic for these parts, and use this model to predict the behaviour of the potential TES implementation. Following on from this we built the TES and characterised it *in-vivo*. We present results from experiments in Section 4.4 which showed that tuning the device adversely affects performance and we suggest potential causes of these effects. In Section 4.5, we characterise the dynamics of protein production by the TES and finally in Section 4.6 we summarise our findings and discuss potential improvements to the design.

4.2 A simple motif for gene regulation

The TES is a simple genetic regulatory motif whose steady state response function can be tuned. The motif has two inputs. A main input regulates transcription of the gene of interest whilst the second tuner input modifies the output relative to the input. By modifying the tuner, the response function of the device can be tuned **Figure 4.1**.

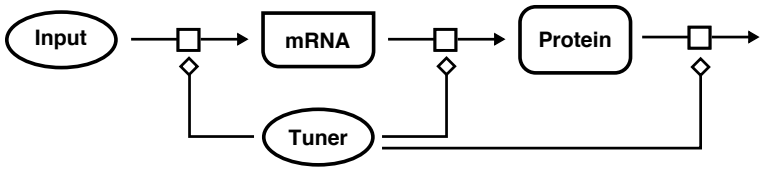


Figure 4.1: Systems Biology Graphic Notation (SBGN) diagrams of tuneable expression systems (TES). SBGN diagram of a TES. A tuner chemical species modifies the reaction producing mRNA, the reaction converting mRNA into protein, or modifies the protein itself in order to tune protein production.

4.2.1 A small RNA regulated tuneable expression system

We developed a general device that implemented the TES motif. The input to the device is the activity of an input promoter that regulates the rate of transcription of a gene of interest, where promoter activity is the flux of RNAP at the promoter. Increasing the input promoter activity increases the rate of transcription of the gene of interest and therefore the concentration of mRNA produced, which in turn increases the rate of translation and therefore the devices output which is the rate of protein production. The relationship between the devices input and output is captured in its response function. By using promoter activity as the device's input, any other device that uses a promoter as its output can be used to regulate the input, allowing the device to be more easily integrated into larger circuits [100].

To tune the response function, the TES uses a second mode of gene regulation to modify the rate of protein production, so that at any input promoter activity, the output can be tuned, allowing us to change the shape of the TESs response function. For example, the rate of induced premature transcription termination [28, 29, 84, 127], translation [16, 17, 54, 55, 68] or protein degradation could be used as a second mode of gene regulation. Whilst we could use many modes of gene regulation to vary the rate of protein production, we focused on designing a device that uses small RNAs (sRNA) to modify the rate of protein production [16, 17, 28, 29, 54, 55, 68, 84, 127]. We chose this class of regulator because: producing sRNAs in cells is not resource intensive, many sRNA devices perform well, some can be designed to respond to endogenous genes in a host, they encompass a wide range of mechanisms for regulating gene expression and many can be predictably designed computationally because RNA thermodynamic models accurately capture their behaviour.

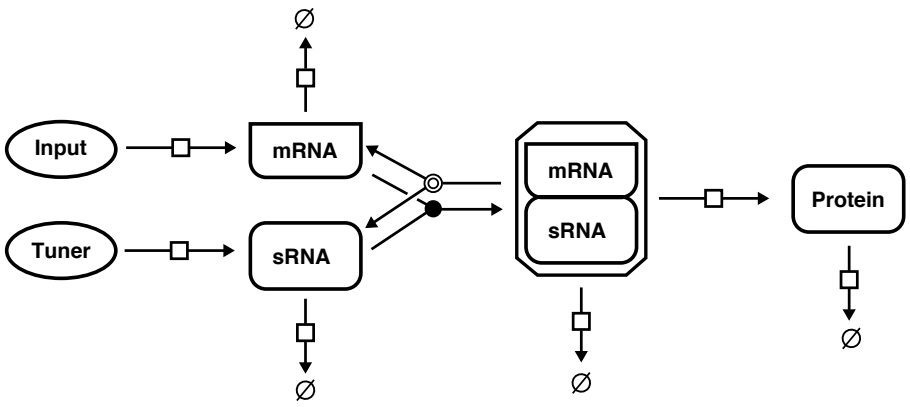


Figure 4.2: Systems Biology Graphic Notation (SBGN) diagram of a sRNA tuneable expression system (TES). SBGN schematic of a TES. Small RNAs (sRNA) modify the rate of protein production. Small squares represent processes, circles with circles inside represent the dissociation of bound molecules, black filled circles represent the association of molecules.

The precise mechanism that sRNA devices use to regulate protein production varies between system. In general, sRNAs interact with mRNA in such a way that alters the effective concentration of mRNA that can be translated, thus modifying the rate of protein production. sRNAs can either interact with mRNAs to increase or decrease the concentration of translatable mRNAs, which results in systems that increase or decrease the rate of protein production in response to increasing rates of sRNA transcription. We focused on sRNA devices that activate protein production in response to sRNAs. The rate of transcription of the sRNA tuner molecule was regulated by the activity of a second promoter – the tuner input **Figure 4.2**.

4.2.2 Modelling a small RNA tuned tuneable expression system

The TES activates protein production in response to sRNAs associating with the mRNA, so toehold switches (THS) [55] or synthetic riboregulators [16, 17, 68] could be used to tune translation. We assume that the sRNA and mRNA associate and dissociate in a reversible process.

Based on reactions between significant biochemical species in the TES we built an ODE model that would allow us to predict how a TES design would behave. We considered reactions involving four biochemical species in the TES:

$$\xrightarrow{r_M} M, \tag{4.1}$$

$$\xrightarrow{r_S} S, \tag{4.2}$$

$$M + S \xrightarrow{k_C^+} C, \tag{4.3}$$

$$C \xrightarrow{k_C^-} M + S, \tag{4.4}$$

$$C \xrightarrow{\beta_c} P + C, \tag{4.5}$$

$$M \xrightarrow{\gamma_M} \varnothing, \tag{4.6}$$

$$S \xrightarrow{\gamma_S} \varnothing, \tag{4.7}$$

$$C \xrightarrow{\gamma_C} \varnothing, \tag{4.8}$$

$$P \xrightarrow{\gamma_P} \varnothing. \tag{4.9}$$

M, S, C and P are the concentrations of : mRNA containing a transcription regulating region of RNA followed by the open reading frame of a gene of interest, sRNA used to tune rate of protein production, complex formed between M & S and protein, respectively. M and S are transcribed at rates of r_M and r_S. mRNA and sRNA bind to form a complex at a rate of $k_C^+ MS$. Complex dissociates into mRNA and sRNA a a rate of $k_C^- C$ and mRNA is translated into protein at a rate of r_P. mRNA, sRNA, complex and protein degrade at a rate proportional to their concentrations and rate constants: γ_M, γ_S, γ_C and γ_P respectively.

Based on this set of chemical reaction equations, the following system of ordinary differential equations were derived:

$$\frac{dM}{dt} = r_M - k_C^+ MS + k_C^- C - \gamma_M M, \tag{4.10}$$

$$\frac{dS}{dt} = r_S - k_C^+ MS + k_C^- C - \gamma_S S, \tag{4.11}$$

$$\frac{dC}{dt} = k_C^+ MS - k_C^- C - \gamma_C, \tag{4.12}$$

$$\frac{dP}{dt} = \beta_C C - \gamma_P. \tag{4.13}$$

4.3 Design and modelling of a tuneable expression system

4.3.1 Design and biological implementation of a tuneable expression system

We designed a specific implementation of the TES that would allow us to test the motif experimentally. We chose parts for the device based on previous characterisation data. The input to our design is a small molecule sensor that regulates the activity of a promoter and a THS is used to regulate translation. The THS forms a hairpin secondary structure that encapsulates both the RBS and start codon of our down stream gene of interest, preventing ribosomes binding to and translating the gene. Translation is activated when a separately transcribed sRNA binds

to the THS. This exposes the RBS and start codon, enabling ribosomes to translate the gene. Translation of a protein regulated by a THS is activated by increasing the concentration of separately transcribed sRNA. We used a second genetic sensor to regulate the rate of sRNA transcription. The TES was used to regulate production of a yellow fluorescent protein (YFP), allowing us to characterise the rate that our device produced protein in single cells by measuring their fluorescence. Our device was assembled as a single plasmid, with which cells were transformed to test the device.

We used a THS to regulate the rate of translation of our gene of interest for several reasons. Firstly, THSs have been shown to strongly repress translation; there is a large fold change between the amount of gene expressed when sRNA is present to activate translation, and when sRNA is absent and translation is repressed. Secondly, 26 THSs have been built to date and shown to regulate different genes within the same circuits. Out of all THSs, when sRNA was transcribed along side a non-cognate THS, translation was only activated to 12 % of its maximum rate [55]. This shows that multiple THSs could be used within the same circuit without interfering with each other. In addition to these 26 devices, a large number of novel THS designs are possible, more so than other synthetic riboregulators. This is because the THS sequence is not limited by the need to include the start codon or Shine-Dalgarno region in their sequence, as is the case with both natural and synthetic riboregulators [16, 17, 55, 68, 98]. By using THSs as a second mode of gene regulation in a TES, if multiple TESs need to be used in the same circuit, new TES designs would not be limited by the availability of new THSs. Finally, THSs could activate translation in response to mRNAs native to a host strain, allowing us to regulate rates of protein production in response to dynamic cellular processes happening in a host, like metabolism or growth. This would make it possible to incorporate the TES into novel circuits that can tune and fix themselves, making them robust to changes in their host's behavior that happen in response to changes in their environment. These properties make THSs perfect for tuning protein production in our TES design, as they would allow us to use many in larger circuits and build versatile circuits applicable to real world conditions.

Whilst a host's native mRNA could have been used to dynamically tune protein production, we used a genetic sensor to produce the synthetic sRNA that activates a THS system (THS) that had already been characterised [55]. We used promoter P_{tac} to regulate the rate of transcription of the translation activating sRNAs. The activity of P_{tac} increases in response to an increase in concentration of small molecule isopropyl β-D-1-thiogalactopyranoside (IPTG). To halt transcription of sRNA we used the terminator L3S3P11 after the sRNA [32] (**Figure 4.3**).

The region of the circuit encoding the sRNA is immediately followed on our circuit plasmid by a region of DNA encoding our gene of interest regulated by a toehold switch that is activated by the upstream sRNA. The activity of promoter P_{tet} was used to regulate the rate of transcription of the THS and downstream YFP. The activity of P_{tet} increases in response to small molecule anhydrotetracycline (aTc).

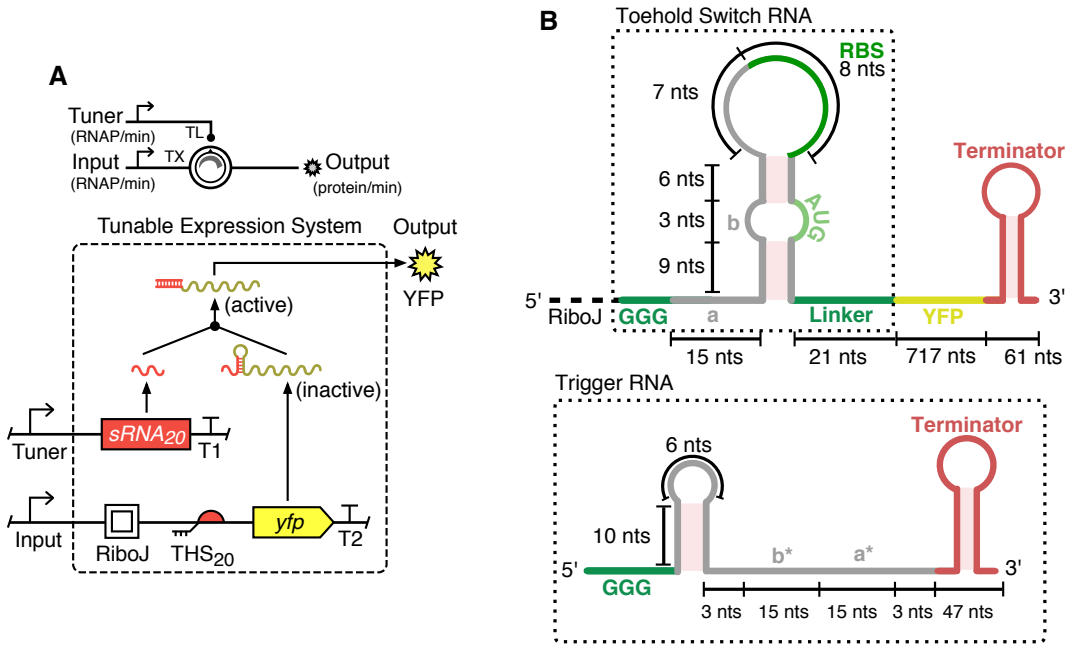


Figure 4.3: Design of a tuneable expression system (TES) (**A**) Sensors regulate rates of transcription (TX) and translation (TL) to control the output from the TES (top). The input regulates transcription of a yellow fluorescent protein (YFP), an upstream toehold switch (THS design 20) [55] renders the transcript inactive to translation until it binds to small RNA 20 ($sRNA_{20}$) and is made active (bottom). T1 and T2 represent transcriptional terminators L3S3P11 and L3S2P21. (**B**) Schematic of toehold switch (THS) implementation. Regions on the THS RNA marked a and b (top) hybridize with their reverse compliments on the trigger RNA (bottom) marked a* and b*, respectively (adapted from [55]).

Genes are transcribed from a transcription start site, which is upstream of a gene's translation start codon (AUG). Only mRNA downstream of the start codon is translated, meaning there is a region of transcribed RNA at the 5'-end of an mRNA transcript that is not translated. This region, known as the 5'-untranslated region (5'-UTR), affects the rate of translation of a gene [83]. Changing a promoter in a transcriptional unit changes the 5' UTR which in turn affects the rate of translation. To insulate our gene of interest from this effect and allow us to change the promoter we used without seeing unpredictable effects on circuit function, we used a RiboJ insulator [83]. RiboJ is a 70 nt sequence that consists of a hammerhead ribozyme at its 5'-end and a hairpin at its 3'-end. Once transcribed, the hammerhead ribozyme cleaves itself, removing the upstream 5'-UTR. Following the insulator, THS number 20 (THS-20) [55] regulates the rate of translation of our gene of interest, YFP. Transcription of YFP is terminated by the L3S2P21 terminator [32] (**Figure 4.3**).

4.3.2 Modelling and simulation of a biological implementation

Before building the TES, we adapted our model found in **Eqs.(4.10 – 4.13)** to predict how the TES might behave. We adapted the model so that it accounts for basal rates of transcription from

the input and tuner sensors. So r_M and r_S become:

$$r_M = \alpha_M^0 + \alpha_M \tag{4.14}$$

and

$$r_S = \alpha_S^0 + \alpha_S \tag{4.15}$$

where α_M^0 and α_S^0 are the basal rates of transcription of mRNA and tuner sRNA and α_M and α_S are the activities of the input and tuner promoters.

We derived biologically relevant parameters (**Table 4.1**) that we used to simulate the TESs behaviour. We found that as the rate of mRNA transcription increased, there was a sigmoidal increase in the number of proteins produced at steady state (**Figure 4.4**). As the rate of tuner transcription increased, the rate of protein production increased at all input levels, transforming the response function (**Figure 4.4**). The increase in the rate of protein production for an input when the tuner promoter activity was increased was non-linear. There is an optimal range of sRNA transcription rates for tuning the response function and shifting the sRNA transcription rate outside of this range has little effect on the rate of protein production.

Within this optimal intermediate range of sRNA transcription rates, which is between 0.012–7.65 RNAP min^{-1}, increasing the rate of sRNA transcription over one interval (from 0.3 to 1.5 RNAP min^{-1}) resulted in a 2.6-fold increase in protein production when the input was 1 RNAP min^{-1} (**Figure 4.4**). This can be seen from **Figure 4.4**, where, within the intermediate range of sRNA transcription rates, the same relative increase in sRNA transcription rate shifts the response function upwards. Outside of this range, increasing the rate of sRNA transcription has little effect on the rate of protein production. Below and above the intermediate range, an equivalent relative increase in rate of sRNA transcription, from 0.0001 to 0.0124 RNAP min^{-1} below the range and from 7.65 to 190 RNAP min^{-1} above the range, yielded only 1.04–fold and 1.3–fold increases in the rate of protein production. This can be seen from **Figure 4.4** where at low and high rates of sRNA transcription, different rates of sRNA transcription yield overlapping response functions that are almost identical.

4.4 In-vivo performance of a tuneable expression system

A genetic circuit implementing the TES was assembled as a plasmid (**Figure 4.5**) and *E. coli* cells were transformed with this construct. We used flow cytometry to measure the fluorescence of populations of these cells when they were in exponential growth, i.e. when protein production was at steady state. To alter the rate of transcription and translation of YFP, populations of cells were grown in different combinations of inducers aTc and IPTG. This was done for three biological replicates. The median fluorescence of populations for each biological replicate was calculated and the mean of these values was found. Next, to find the input and tuner input promoter activities

Table 4.1: Parameters used to model tuneable expression systems

Name	Description	Value(s)	Unit	Ref.
α_M	Induced transcription rate of input promoter	0 to 300[a]	RNAP min^{-1}	[124], [72], [61]
α_S	Induced transcription rate of tuner sRNA promoter	0 to 190[b]	RNAP min^{-1}	[124], [72], [61]
α_M^0	Basal transcription rate of input promoter	0.0886[c]	RNAP min^{-1}	[124], [72], [61]
α_S^0	Basal transcription rate of tuner sRNA promoter	0.2307[c]	RNAP min^{-1}	[124], [72], [61]
β_P	Rate of protein production from active toehold switch	5	Proteins $complex^{-1}$ min^{-1}	[105]
k_C^+	Association rate of sRNA and mRNA into complex	0.0257[d]	complexes $transcript^{-1}$ min^{-1}	[137]
k_C^-	Dissociation rate of complex into sRNA and mRNA	0.00672[d]	transcripts $complex^{-1}$ min^{-1}	[137]
γ_M	Degradation rate of mRNA	0.231[e]	min^{-1}	[7]
γ_S	Degradation rate of sRNA	0.231[e]	min^{-1}	[7]
γ_C	Degradation rate of mRNA:sRNA complex	0.231[e]	min^{-1}	[7]
γ_P	Degradation rate of output protein	0.035[f]	min^{-1}	[81]

a. We assume the maximum rate of transcription initiation of P_{tet} is the same as a constitutively expressed gene [72], and that a plasmid with a p15A origin of replication has 15 copies per cell [61].
b. The maximum transcription initiation rate of Ptet is 1.57 times higher than Ptac.
c. Maximum rates of transcription initiation of sensors, α_M and α_S, were divided by fold changes for P_{tet} and Ptac as measured in previous work [124].
d. Assuming rates of RNA hybridization are similar to DNA hybridization [137].
e. We assume that the half-life of the THS transcript and tuner sRNA are similar to the average mRNA half-life in a cell, measured to be 3 minutes in exponentially growing cells [7].
f. We assume dilution of proteins caused by cell growth is the main cause of degradation and cell doubling time of 20 min.

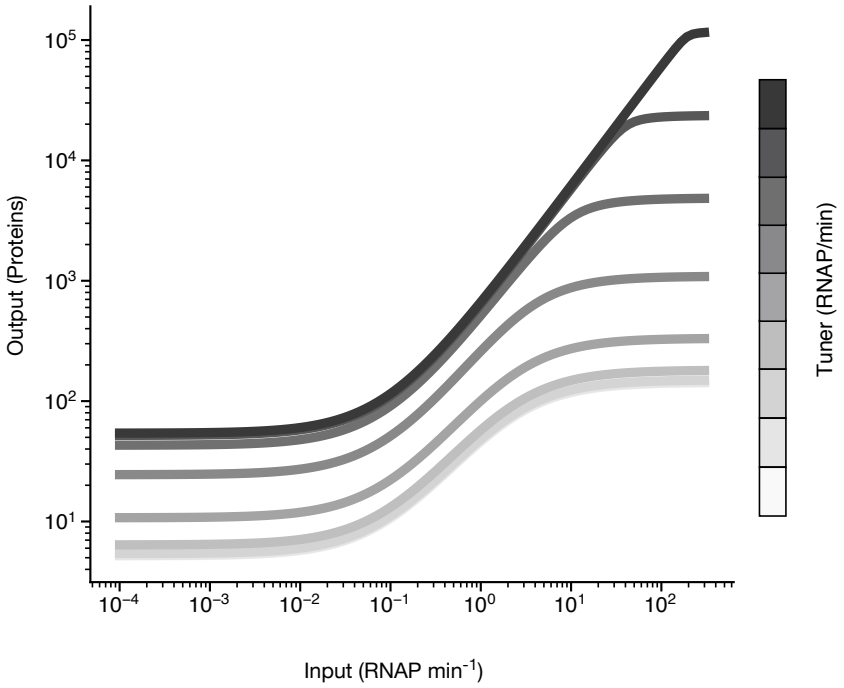


Figure 4.4: Modelled response function of tuneable expression system (TES). Response functions show number of proteins at steady state found using our deterministic model of the TES, presented in **Eqs. (4.10 – 4.13)**, simulated at varying tuner inputs. Line colour indicates promoter activity of tuner inputs, from light to dark (0.0001, 0.0005, 0.0025, 0.012, 0.062, 0.31, 1.53, 7.6, 38.0, 190.0 RNAP min-1).

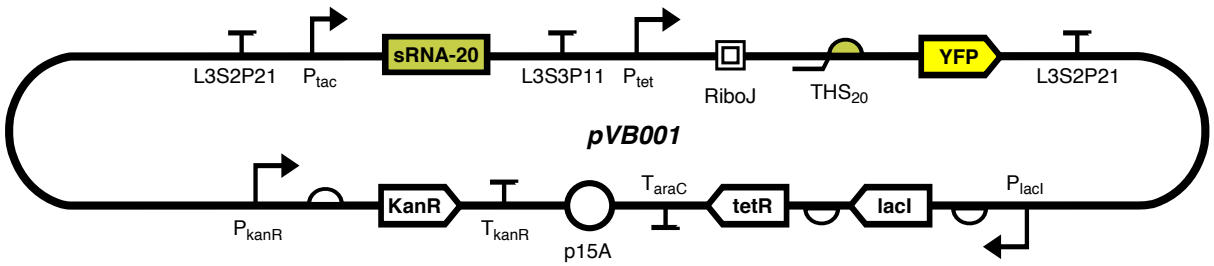


Figure 4.5: Schematic diagram of tuneable expression system testing plasmid. Synthetic Biology Open Language visual (SBOLv) glyphs are are used to demonstrate our plasmid design [87] *KanR* is the kanamycin resistance gene, p15A is the origin or replication. *tetR* is the coding sequence for the tetracycline repressor protein, the transcription factor used in the P_{tet} sensor system. lacI is the coding sequence for the lac inhibitor protein, the transcription factor used in the P_{tac} sensor system.

(outputs from P_{tet} and P_{tac}), relative to the concentrations of aTc and IPTG, the inputs to P_{tet} and P_{tac}, respectively we characterised the sensors on their own.

4.4.1 Characterising sensors to find inputs to the tuneable expression system

To find the relationship between the input promoter activities and the inducer concentrations that are the inputs to the sensors we transformed cells with circuit characterisation plasmids. The plasmids each contained one of the sensors which was used to drive the expression of YFP. We grew these cells in a range of inducer concentrations and measured their fluorescence using flow cytometry. We found the response function for the sensors (**Figure 4.6**), which relates inducer concentration to promoter activity in terms of cell fluorescence. We converted the sensor outputs from cell fluorescence in arbitrary units (a.u.) into relative promoter units (RPU), a measure of promoter activity calculated by dividing a fluorescence measurement by the fluorescence given off by a standard genetic construct that constitutively expresses YFP (Chapter 3, Section 3.4.5) [19, 100]. By converting the sensor outputs to RPUs, we can compare our device performance to other devices [100]. Both promoters are leaky, that is, they're active when no inducer is present. In the absence of inducer, we see output promoter activities for P_{tet} and P_{tac} of 0.002 and 0.003 RPU, respectively. We also see a significant difference in the maximum outputs of each device, with P_{tet} giving an output of 6.6 RPU, which is 2.4-times higher than P_{tac} (2.75 RPU).

4.4.2 Response to increased transcription rates

We characterised the TES over a range of different combinations of aTc and IPTG concentrations. By varying concentrations of aTc and IPTG we changed the activity of promoters P_{tet} and P_{tac}, respectively, which varied the rates of transcription of mRNA and sRNA. As we increased the rate of mRNA transcription, we saw a sigmoidal increase in the amount of protein produced. We see this for all tuner promoter activities we tested (**Figure 4.7B**). There was little increase in output as the input was increased from 0.002 to 0.5 RPUs. As the input increased above 0.5 RPU, the amount of protein produced increased when the input promoter activity was increased (**Figure 4.7B**). This is as we expected from our model (**Figure 4.7A**).

At rates of sRNA transcription less than 0.03 RPU, as we increased the rate of transcription of mRNA, we saw a 14-fold increase in the rate of protein production (**Figure 4.7**). The THS we used has been shown to almost fully repress translation of a gene until the sRNA is added [55], so when rates of mRNA and sRNA transcription were lowest, we expected no protein to be produced. However, we measured 25.6 a.u. of fluorescence. This could be due to: the THS not functioning properly, leaky transcription from the tuner promoter P_{tac} causing sRNA to be produced which unfolds the THS and activates YFP production, or a combination of these. At its lowest output activity, the tuner promoter is still active at a level of 0.003 RPUs, suggesting that sRNA is still being produced. This could explain why the output from the TES increases with an increase in the rate of mRNA transcription when the tuner promoter activity is low.

Fold-change is a useful characteristic of a device's performance [130]. We calculated the fold-change of the TES for low and high outputs at all tuner inputs and found it decreased significantly

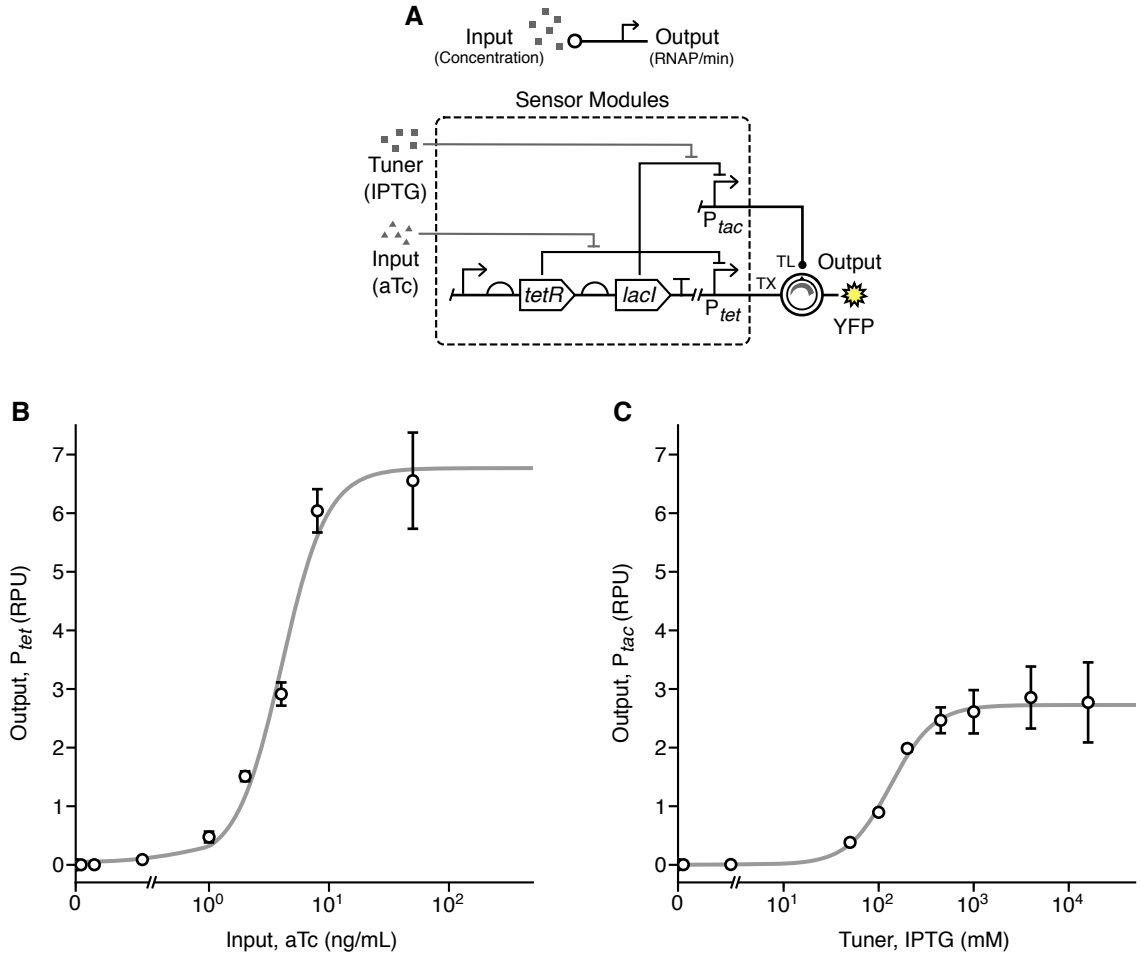


Figure 4.6: Characterisation of sensors. (A) Design of sensor modules. An increase in the concentration of sensor input increases the rate of transcription of the sensor (top). aTc increases activity of input promoter, P_{tet}, whilst IPTG increases activity of tuner promoter, P_{tac} (bottom). **(B)** Response function for P_{tet} **(C)** Response function for P_{tac}.In both response functions, points are means for 3 biological replicates, errorbars show ±1 standard deviation. Output promoter activities are given in relative promoter units (RPU).

as we increased the tuner promoter activity (**Figure 4.8A**). Fold change dropped from 14.1 at low tuner inputs (P_{tac} = 0.003 RPU), to 2.4 at high tuner inputs (P_{tac} = 2.61 RPU). Whilst the fold-change decreased with an increase in tuner input, the absolute difference between the output when the rate of mRNA transcription is highest and lowest, the dynamic range, increased from 53 at low tuner promoter activity (P_{tac} = 0.003 RPU) to 695 at high tuner promoter activity (P_{tac} = 2.61 RPU) (**Figure 4.8B**). The fold change decreased despite an increase in dynamic range because the the rate that protein was produced by the TES at the lowest input promoter activity (y_{min}), increased significantly with an increase in the rate of sRNA transcription (**Figure 4.8C**).

The parts we used to build the TES define the maximum and minimum rate that the TES can produce protein. Based on how the input promoter (P_{tet}), tuner promoter (P_{tac}) and THS

behave individually, we expected to see higher fold changes in output. Assuming mRNA is only produced due to P_{tet} activity, the number of active mRNAs able to be translated will be less than the total number of mRNAs transcribed due to P_{tet} activity, because, due to the THS, only a fraction of mRNA transcribed (those bound by sRNA) are translated. We assume that when the tuner promoter activity is highest (P_{tac} = 2.61 RPU), at the lowest input promoter activity (P_{tet} = 0.002 RPU), there is an excess of sRNA compared to mRNA, so the majority of mRNAs are bound and can be translated. Under these inputs, the number of active mRNAs bound to sRNAs, is similar to the number of mRNAs transcribed by P_{tet} at its lowest activity. This rate of protein production, $y_{min'}$, is the amount of active mRNA produced by the TES when the input promoter activity is lowest (P_{tet} = 0.002 RPU), and the activity of the tuner promoter is highest, (P_{tac} = 2.61 RPU).

The highest rate that the TES can transcribe mRNAs is less than or equal to the highest rate that RNA can be transcribed by the input promoter when it is most active (when P_{tet} = 6.6 RPU). mRNA is translated when it is bound to sRNA, so the number of mRNAs that are active and can be translated is limited. This further limits the maximum rate that the TES can produce protein. Assuming all sRNA molecules bind to and activate mRNAs, the number of active mRNAs available to be translated is the same as the number of sRNAs produced when the output from the tuner input is most active (P_{tac} = 2.61 RPU). The TES produces protein

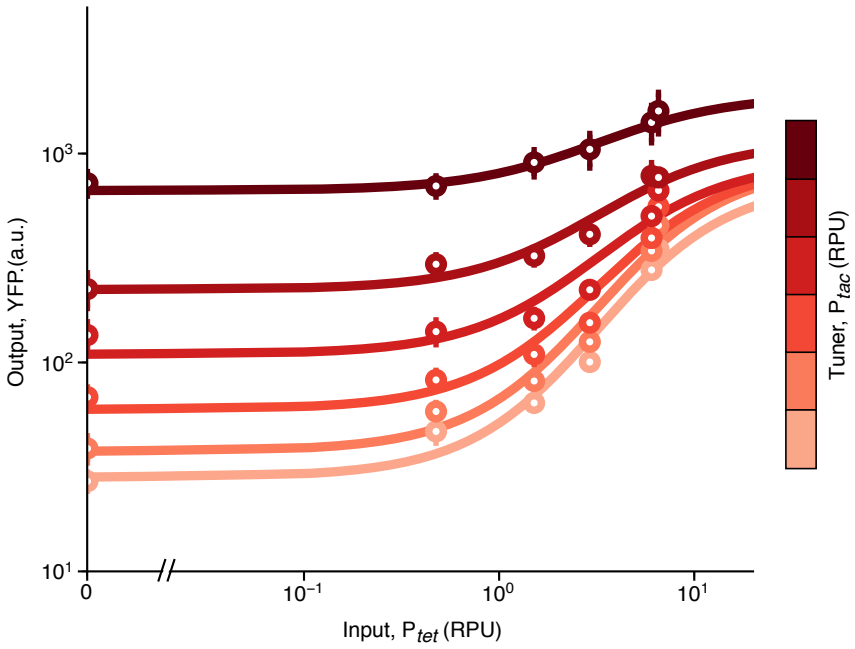


Figure 4.7: Response of the tuneable expression system (TES) to changes in input and tuner promoter activities. Experimental response function of the TES, points are the mean of three biological replicates and error bars show standard deviation. Lines are fitted hill functions at different tuner inputs (light–dark: 0.003, 0.03, 0.15, 0.43, 0.9, 2.6 RPU).

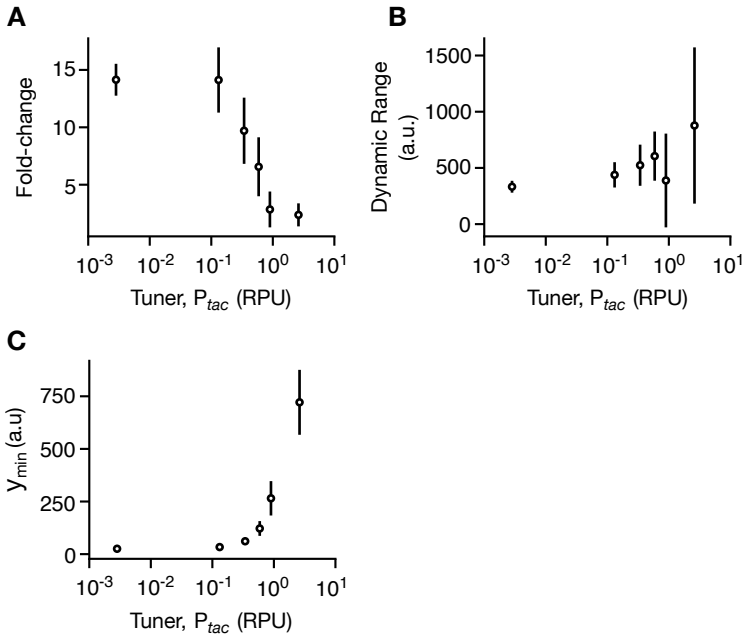


Figure 4.8: Effect of tuner input on device performance. (**A**) Fold change of the TES against tuner promoter activity. (**B**) Dynamic range against tuner promoter activity (**C**) Minimum protein production in response function (y_{min}) against tuner promoter activity. Points are means calculated from three biological replicates, error bars are ± 1 standard deviation.

at maximum rate, Y_{TES}^{max}, when input and tuner promoters, P_{tet} and P_{tac}, are fully active and their outputs are 6.6 RPU and 2.61 RPU, respectively. Under this configuration of inputs, the number of mRNA molecules bound to sRNA that can be translated is the same as the number of sRNA molecules that P_{tac} produces when it is most active. Assuming the sensors produce the same amount of RNA in the TES as they did when they're characterised independently, we can estimate the fold change between outputs from the TES at lowest and highest input promoter activities (corresponding to 0.002 RPU and 6.6 RPU, respectively) when the tuner input is most active (P_{tac} = 2.61 RPU).

Under these conditions, we would estimate a fold-change of 1305. However, what we see from an equivalent configuration of inputs to the TES is a fold-change of 2, four orders of magnitude lower than we expect.

A TES will have a low fold-change if the rate it produces protein at is too high at low input promoter activities or too low when the input promoter activity is high. The TES's output when the input promoter is most active (P_{tet} = 6.6 RPU), could be too low for several reasons. First, the TES's input and tuner promoters might not function in the TES in the same way as they did when they were characterised independently. P_{tet} and P_{tac} might be less active when they're used in the TES than when they were characterised. There are a couple of key differences in how we use the THS compared to the context in which it was originally used [55]. The first difference is

that we regulate production of sRNA and THS using P_{tac} and P_{tet}. When the THS was originally characterised, sRNA and THS were both regulated by P_{T7} promoters on separate plasmids. sRNA was expressed from a higher copy number plasmid than mRNA, and so was produced in excess of THS. Second, we use a RiboJ insulator upstream of the THS [124]. The RiboJ insulator folds into a strong hairpin structure that could have interfered with the THS. If the folded THS structure is altered around the toehold that the sRNA binds to, it could have reduced the rate that sRNA bound to the THS. This would have in turn reduced the rate of translation, causing the device to produce less protein when both sRNA and mRNA are transcribed at their highest rates.

There are several reasons that the output from the TES could be higher than expected when the input promoter activity is low. Firstly, if the toehold switch does not fold correctly, less energy would be required to unfold the structure and allow ribosomes to translate the mRNA. The rate of translation would increase by a rate proportional to the number of unbound mRNAs causing the number of proteins produced at all inputs to increase. Another explanation for the high concentrations of protein produced is that the terminator after the sRNA is not efficient. If this is the case, some of the RNAPs transcribing sRNA may read through the terminator and cause transcription of the mRNA, even though the input promoter was inactive.

4.4.3 Single cell variability

Gene expression is governed by interactions between biochemical molecules, such as DNA, which can have low copy numbers in cells. Because of this, small fluctuations in the copy number of a molecule can have a significant impact on the total number of copies in a cell. The copy number of molecules involved in gene expression (i.e. the number of DNAs, mRNAs, sRNAs and proteins) varies significantly over time and between different cells, making gene expression a stochastic process [41, 110].

As gene expression is stochastic, the concentrations of an expressed protein will vary in individual cells growing in a population. For a genetic device to propagate a signal, its output must change sufficiently when its input promoter activity switches. If a device's input promoter activity is low, but due to stochastic gene expression, a large percentage of cells harbouring the device have outputs that suggest the input promoter activity is high, this will be problematic.

Based on average data for populations of cells harbouring the TES, when the tuner input promoter activity is increased from low (0.003 RPU) to high (2.61 RPU), the fold-change of the device decreased from 14 to 2.1. This suggests that, as the input promoter activity was increased from low (0.002 RPU) to high (6.6 RPU) the amount of protein produced by cells in a population varied significantly (**Figure 4.7**). However, when we looked at distributions that show the concentration of protein in individual cells in a population, we saw that, the distributions overlapped as input promoter activity was increased from 0.002 RPU to 6.6 RPU (**Figure 4.9**). For a large percentage of cells it is difficult to determine, based on the amount of protein the TES produces in the cell, if the device's input promoter activity is low or high.

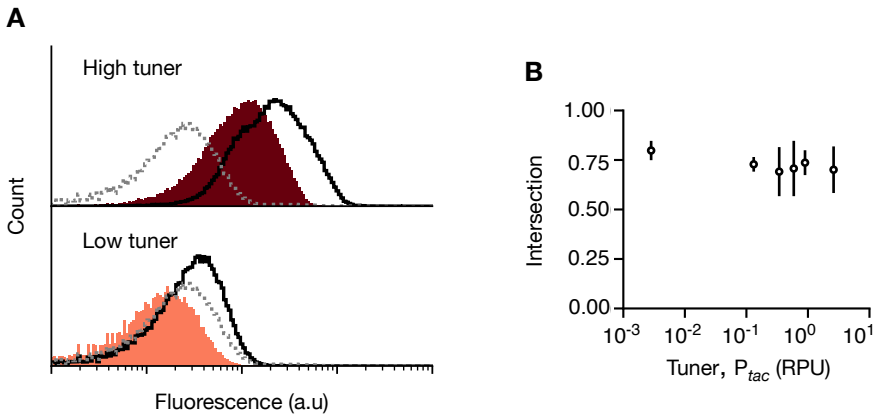


Figure 4.9: Population wide performance of the tuneable expression system. (**A**) Distributions of YFP fluorescence in populations of cells when the tuner promoter activity is low (bottom: 0.003 RPU) and high (top: 2.61 RPU). Filled and black outlined distributions are for cells grown with low (0.002 RPU) and higher (6.6 RPU) promoter activities, respectively. The autofluorescence distribution has a grey dotted line. (**B**) Intersection between fluorescence distributions when input promoter activity if high and low. Points are means calculated from three biological replicates, error bars are standard deviations.

We measured the overlap between population distributions when the input promoter activity was high (6.6 RPU) and low (0.002 RPU) as a percentage. This measurement is the intersection between distributions [126] and at all tuner input promoter activities we measured, it was high (> 69%) (**Figure 4.9B**). Given two populations of cells harbouring the TES grown at high and low inputs, it would be difficult to tell if any one individual cell from either population had an input that was high or low.

4.5 Input promoter activities affect dynamics of protein production

We wanted to see how changing the TES's input and tuner promoter activities changed the rate of protein production over time. The rate of protein production is related to the rate of cell growth so we first determined whether growing cells in different inducer concentrations affected the growth rate of cells harbouring the TES.

4.5.1 aTc and IPTG concentrations do not affect cell growth

To measure how the rate that the TES produces protein changes with time we grew cells harbouring the TES in varying concentrations of aTc and IPTG. At regular intervals, a robotic arm transferred the plate of cells to a plate reader that measured the optical density (OD_{600}) and fluorescence of populations of cells. Looking at the growth curves we generated, we see that the two different biological replicates we tested do not grow in sync. For one of the replicates, there is

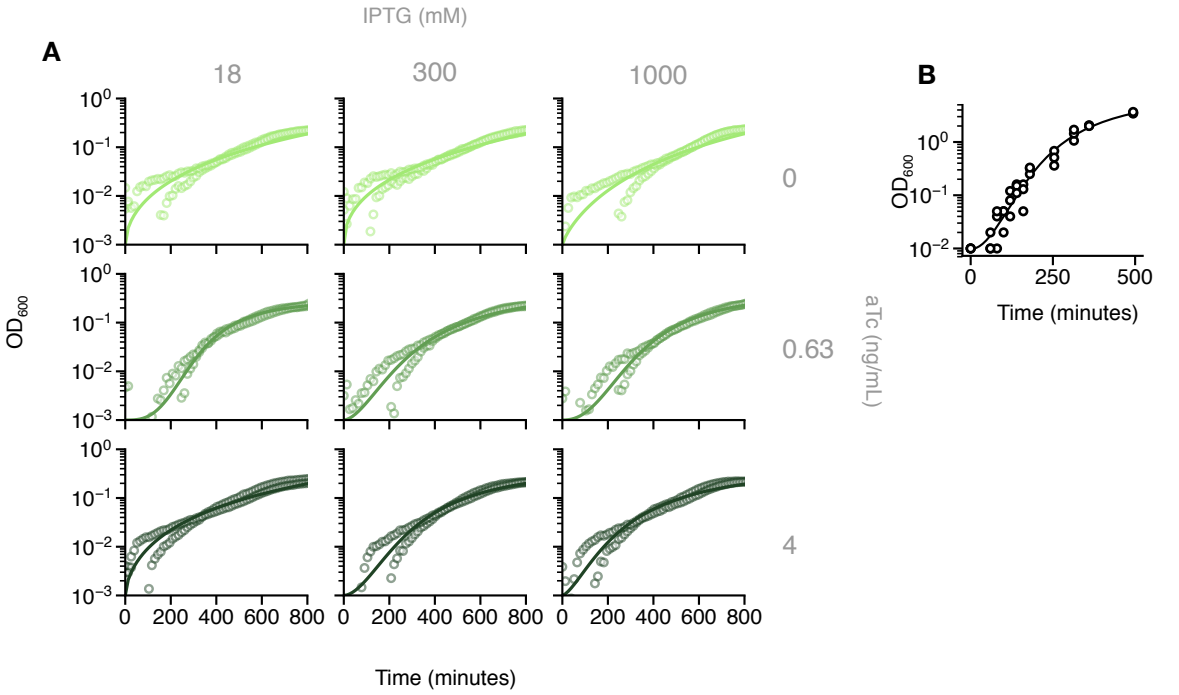


Figure 4.10: Growth of cells containing a tuneable expression system. (**A**) Growth curves for populations of cells harbouring the tuneable expression system (TES) grown in varying concentrations of inducers aTc and IPTG. Cultures were inoculated with cells growing in stationary phase. We assume that optical density (OD_{600}) is proportional to cell density. All points denote OD_{600} for one of two biological replicates. Growth curves at all inducer concentrations can be found in **Figure A.1**. Solid lines denote growth curve fit to a mathematical model [35]. (**B**) Growth curve for cells harbouring a tuneable expression system grown with no inducers, in media inoculated with cells growing in exponential phase. Points denote OD_{600} for one of three biological replicates. Solid lines denote growth curve fit to a mathematical model [35]

a lag of 200 minutes before cells start growing exponentially, for the other, cells start growing exponentially at the start of growth (**Figure 4.10A**).

The growth of cells is not consistent across the two replicates. Growth curves for populations of cells grown in the initial (steady state) set of experiments we ran, showed that across biological replicates there was little variation in cell growth (**Figure 4.10B**). In these experiments, the growth curves of all the replicates were well synchronized. The difference between these two sets of experiments was that in the initial experiments, we inoculated media with cells that were already in their exponential growth phase. In the dynamic experiments, media was inoculated with cells in stationary phase (**Figure 4.10A**) (see: Chapter 3, Sections 3.4.3 and 3.4.1 for cultures inoculated with cells growing in steady state and exponential phases, respectively). When characterising circuits, to get reproducible results, media must be inoculated with cell populations already growing at exponential phase. When media is inoculated with cells growing in stationary phase, there could be a lag phase before cells start growing and as we show here, this will happen in some populations and not others, so growth was inconsistent across biological

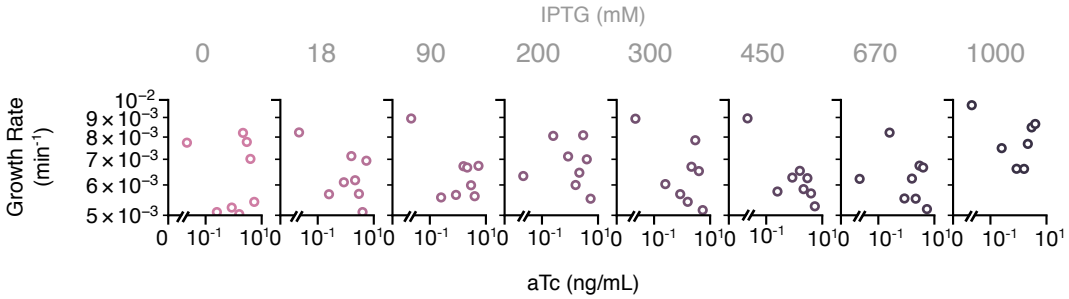


Figure 4.11: Effect of inducer concentration on growth rate. Growth rate is the maximum gradient of the curve that was fit to growth data (**Figure 4.10**).

replicates.

We fitted growth curves for populations of cells with a mathematical model (see: Chapter 3, Materials and Methods, Section 3.4.8). We used numerical methods to find the steepest gradients in the growth curves, which are the growth rates of the cell cultures. aTc and IPTG concentrations did not affect growth rate, showing that using these inducers and sensor systems did not have a toxic effect on cell growth (**Figure 4.11**).

We used a plate reader to measure the change in the average fluorescence of populations of cells harbouring the TES, grown in combinations of aTc and IPTG concentrations, once every 16 minutes (Section 3.4.3). A particle swarm optimisation algorithm was used to fit our model to the data (Chapter 3, Section 3.5.2) [71]. We related the input and tuner input promoter activities (α_m and α_s, respectively in **Eq. (4.10 - 4.13)** to inducer concentrations using a standard hill equation of the form:

$$\alpha = C\left(y_{min} + (y_{max} - y_{min})\frac{x^n}{K^n + x^n}\right), \tag{4.16}$$

where: y_{min} and y_{max} are the minimum and maximum sensor outputs, respectively, n is the hill coefficient, C is a scaling factor that relates fluorescence to RNAP flux and K is the concentration of inducer that induces an output sensor activity to halfway between y_{min} and y_{max}. At higher concentrations of aTc (0.63 and 4 ng/mL), the model fitted experimental data well. However, when there was no aTc, the model failed to capture the TES's behaviour (**Figure 4.12**). Parameters from the optimisation can be found in **Table 4.2**. Looking at the experimental data there is clearly a problem with how fluorescence was measured, as we do not capture any fluorescence below around 100 a.u. (**Figure 4.12**). This is likely due to the sensitivity of the instrument we used to measure fluorescence (Chapter 3, Section. 3.4.3), or the parameters used to set up the instrument.

We plotted the time that fluorescence was detectable (i.e. over 100 a.u.), against aTc concentration and IPTG concentration (**Figure 4.13**). Increasing the concentration of aTc decreased the time taken for cells harbouring the TES to produce protein at a rate that made fluorescence in the cells detectable. Increasing the concentration of IPTG further decreased the response time of the

Table 4.2: Parameters fitted to experimental data

Name	Description	Value(s)	Unit
y_{min}^{tet}	minimum promoter activity for input promoter P_{tet}	1326	a.u.
y_{max}^{tet}	maximum promoter activity for input promoter P_{tet}	7356	a.u.
n^{tet}	hill coefficient for input promoter P_{tet}	3.5	-
K^{tet}	Inducer concentration that promoter activity for input promoter P_{tet} is halfway between its minimum and maximum	1.26	-
C^{tet}	conversion factor for input promoter P_{tet}	0.0063	RNAP a.u. $^{-1}$ min^{-1}
y_{min}^{tac}	minimum promoter activity for tuner promoter P_{tac}	464	a.u.
y_{max}^{tac}	maximum promoter activity for tuner promoter P_{tac}	7356	a.u.
n^{tac}	hill coefficient for tuner promoter P_{tac}	0.8	-
K^{tac}	Inducer concentration that promoter activity for tuner promoter P_{tac} is halfway between its minimum and maximum	1.2	-
C^{tac}	conversion factor for tuner promoter P_{tac}	0.0038	RNAP $a.u^{-1}$ min^{-1}
β_P	Rate of protein production from active toehold switch	0.51	proteins $complex^{-1}$ min^{-1}
k_C^+	Association rate of sRNA and mRNA into complex	1.33	complexes $transcript^{-1}$ min^{-1}
k_C^-	Dissociation rate of complex into sRNA and mRNA	0.0084	transcripts $complex^{-1}$ min^{-1}
γ_M	Degradation rate of mRNA	0.075[e]	min^{-1}
γ_S	Degradation rate of sRNA	0.446	min^{-1}
γ_C	Degradation rate of mRNA:sRNA complex	0.22 [e]	min^{-1}
γ_P	Degradation rate of output protein	0.0056[f]	min^{-1}

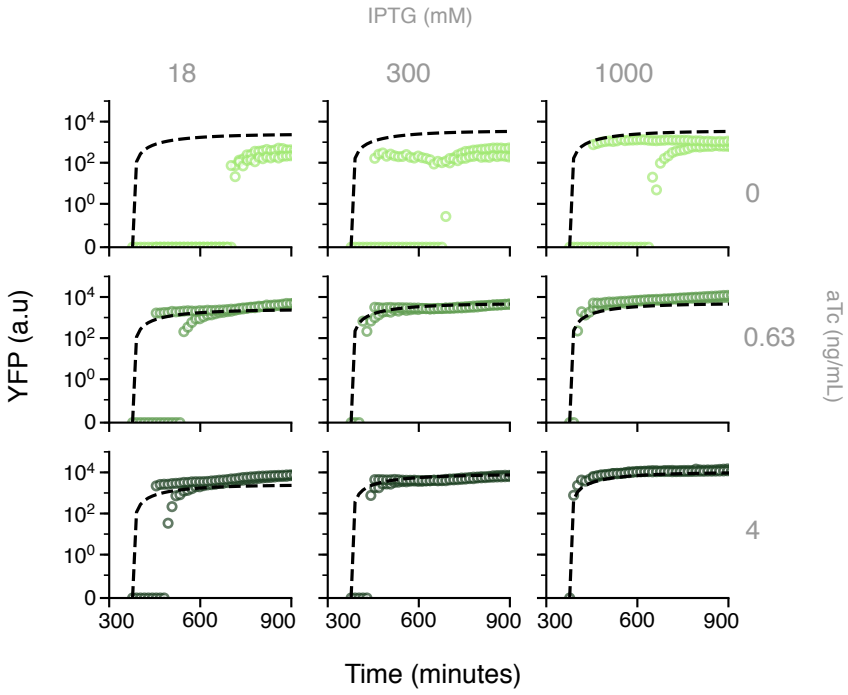


Figure 4.12: Change in the rate protein production over time. Points are for single biological replicates. They represent the average concentration of yellow fluorescent protein (YFP), in arbitrary units (a.u) for a bulk of cells. Change in YFP concentration with time is shown for cells grown in varying concentrations of aTc, which increases the input promoter activity and IPTG, which increases the tuner promoter activity. Dashed black line shows the output from the model, **Eq. (4.10 - 4.13)**, fit to the data. Full data sets can be seen in **Figure A.2**.

device (**Figure 4.13**). At a low tuner promoter activity, associated with an IPTG concentration of 0 mM, increasing the input promoter activity from low to high (aTc concentrations from 0 to 4 ng/mL) made the response 4 times faster. At a high tuner promoter activity, associated with an IPTG concentration of 1000 mM, the same change in input promoter activity results in a 175-fold faster response time. As we expected, increasing the tuner promoter activity and input promoter activity made the TES respond faster.

4.6 Discussion

In this chapter we presented a simple motif for gene regulation which we called a tuneable expression system (TES). The TES uses a promoter as its input which regulates the rate of transcription of a gene of interest. A second mode of gene regulation was used to tune the rate of protein production for any given input. As a proof of concept, we designed a TES that uses toehold switches (THS) to tune production of a yellow fluorescent protein (YFP). The input to the device is the output from a sensor that we used to regulate the rate that YFP was transcribed in response to the small molecule aTc. Upstream of the YFP, a toehold switch folds in a way that prevents

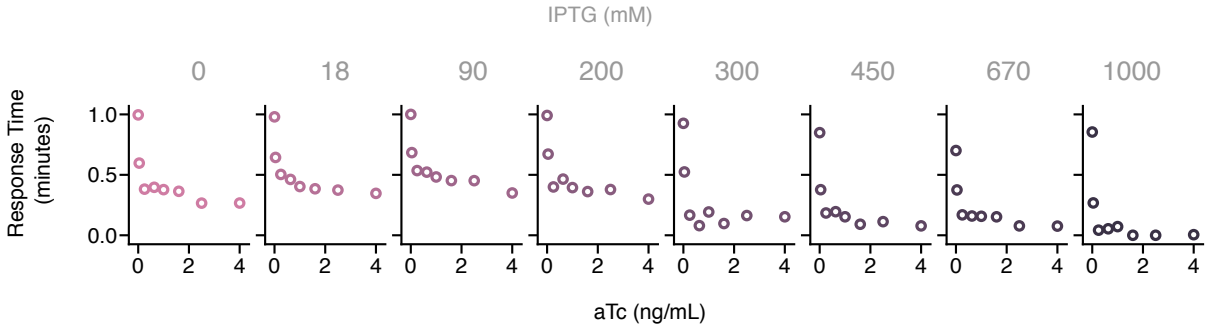


Figure 4.13: Tuneable expression system produces protein at different rates at different inputs. The response time is the time taken for a population of cells harbouring the tuneable expression system to produce an average fluorscence greater than 100 a.u.

ribosomes binding to it, repressing translation. The output from a second sensor increases the rate that a small RNA (sRNA) is transcribed. The sRNA interacts with the THS upstream of YFP to increase the rate that YFP is translated. We built a model that used parameters realistic to the parts in our design and used it to predict how the proposed design might behave.

Next, we physically built and experimentally tested the TES and saw that increasing the rate of sRNA transcription increased the amount of protein produced at all input promoter activities. However, we noticed two issues in the device's performance. First, the fold-change of the TES was lower than we expected; it was 14 when tuner input was high (2.61 RPU) and 2 when it was low (0.003 RPU). Second, for populations of cells containing the TES growing at low and high input promoter activities, at least 69% of individual cells produced protein at the same rate. The percentage of cells in a pair of distributions that produce protein at the same rate is the intersection between the distributions.

We propose a number of reasons for the fold-change of the TES being lower than expected. First, the THS might not have been folding properly, so ribosomes were able to bind to the RBS and translate YFP. In this case, the protein production rate when the inputs were at a minimum would be higher than expected, causing a low fold-change. Alternatively, the THS may form too tight a structure in the conditions we use it in. This would mean that sRNA added to the system doesn't unfold the THS to expose the RBS to ribosomes, so the protein production rate when both input promoter activities are at their highest would be lower than expected. The THS might not have been folding correctly because the RiboJ insulator directly before the THS may interfere with its folding. As is the case with THSs, how a RiboJ works is heavily determined by the secondary structure of RNA that it folds into. It is possible that the RiboJ and THS each interferes with how the other part folds and so affects the performance of their neighbour. If the RiboJ is interfering with the performance of the THS, we can improve the fold-change of the TES by removing it. Increasing the fold-change in this way would also improve the intersection between population distributions.

Similarly, we want to lower the intersection between two distributions of cells when the input promoter activities are high and low. We can do this in one or more of three ways. First, we can reduce the variation in the number of proteins produced per cell in populations of cells. Second, we can increase the average amount of protein produced by cells when the input promoter activity is high. Third, we can lower the average amount of protein produced by cells when the input promoter activity is low. By increasing the rate of sRNA transcription, we can increase the amount of protein produced in cells. This would decrease the intersection whilst also increasing the amount of protein produced. In the design presented here, when the input promoter activity is high, there is always a lower concentration of sRNA produced than mRNA because we produced sRNA using a promoter that has a lower maximum activity than mRNA. We could increase the fold-change of the device by increasing the maximum amount of sRNA that we produce and in doing so, reduce the intersection between populations of cells grown with low and high input promoter activities.

We characterised how the growth of cells harbouring the TES changes depending on input and tuner promoter activities and found that there was no correlation between input and tuner promoter activities and growth rate. We measured how the rate of protein production in cells harbouring the TES changed with time. We used this data to fit our mathematical model using a particle swarm optimization algorithm [71]. We found that at high input and tuner promoter activities, the fitted parameters matched well with experimental data, however, the modelled responses do not match experimental data at low input and tuner promoter activities. As with the growth data, the time course data for protein production (i.e. fluorescence) that we used to fit the model was only made up from two biological replicates. The fitted parameters might fit the data better if more repeat experiments were carried out. We looked at how long it took for populations of cells harbouring the TES to produce enough fluorescent protein for us to detect when the cells were grown in different inducer concentrations. We found that as we increased the input promoter activity by increasing the concentration of aTc, the cells harbouring the TES produced protein faster. When we increase the tuner promoter activity by increasing the concentration of IPTG that cells were grown in, cells produced protein even faster.

The results in this chapter suggest that the TES's performance is limited by two features of its design. First, when characterised alone, the maximum rate that RNA was transcribed by the input promoter (P_{tet}) was more than double that of the tuner promoter (P_{tac}). This suggests that when both promoters were at their most active a lower concentration of sRNA relative to mRNA was produced, resulting in only a fraction of transcribed mRNAs being activated and translated. This would limit the maximum output of the device. Second, the THS we use to regulate translation of the output protein might not be folding correctly due to the adjacent RiboJ insulator. This might have prevented it from sufficiently repressing translation or being activated by sRNAs. If this is the case, we can optimize the TES's performance by boosting the rate of sRNA transcription and removing the RiboJ insulator.

In the following chapter we use a combination of deterministic and stochastic modelling to test if increasing the rate of sRNA transcription will improve device performance. We then use thermodynamic modelling to investigate the effect that the RiboJ insulator has on the THS. Results from modelling are used to inform the design and construction of new TES variants which we then validate experimentally.

CHAPTER 5

Model Guided Optimization of Tuneable Expression System

5.1 Introduction

In the last chapter we built and tested a device called a tuneable expression system (TES) that allows us to tune the rates of transcription and translation of a gene of interest. The device's input is the rate that it transcribes a gene, its output is the rate that it produces protein. The relationship between the two is tuned by controlling the rate the gene is translated, but doing so came at a cost. As we tuned the device the fold-change between its on and off output states significantly decreased, from 14 at low to 2.4 at high rates of translation. We also found that, for two populations of cells grown at high and low inputs, a high number of individual cells in one population produced the same amount of YFP as cells in the other. Under all the conditions we tested, the fraction of the populations that overlapped, the intersection between distributions, was greater than 69%. Based on the these findings, we suggested two causes of these problems.

Firstly, if the maximum amount of protein the TES produces is low, the fold-change of the device will be low. The TES might not produce enough sRNA to enable all the mRNAs we transcribe to be translated, so the amount of protein the TES produces is lower than it could be. This could have caused the device's fold-change to be low. Secondly, we used a self-cleaving ribozyme, RiboJ, to insulate the expression of our gene of interest from the contextual effects caused by its upstream promoter. Once transcribed, the RiboJ insulator folds and self-cleaves, leaving a strong hairpin [34]; this hairpin structure might interfere with the speed that the THS folds and the stability of its final structure. This in turn may have affected the rate of protein production.

In this chapter we use mathematical and biophysical models to optimize the devices design. We then implement and test these improvements *in-vivo*. Through this process we reveal the

causes of the previous performance issues and find novel insights that synthetic biologists can use to build better genetic circuits using RNA–based control elements.

5.2 Design improvements

When the original THS was first developed [55], the THS and sRNA were transcribed using the same promoter but on separate low and high copy number plasmids, respectively. This ensured that the sRNA's production rate was higher than the THS's. The excess of sRNA this produced resulted in most of the THS molecules being bound to sRNAs, enabling them to be translated. In our original TES design, the sRNA and THS were expressed from the same plasmid. sRNA transcription was regulated by P_{tac} and transcription of the THS was regulated by P_{tet}. When we characterised the promoters independently the maximum activity of P_{tet} was more than double that of P_{tac}. We hypothesised that when both promoters were at their highest activities there was an excess of THSs over sRNAs which limited the achievable rate of translation.

To increase the maximum rate of protein production from the TES and potentially improve the fold-change of the device, we could increase the rate that sRNA is transcribed at. The increase this will cause in amount of protein produced by the TES could also decrease the intersection between population distributions for on and off states. Another way we could reduce intersection would be to reduce the variation in the amount of protein produced by cells in a population. Gene expression in cells is a stochastic process, so we expect to see variation in the amount of protein that individual cells produce across a population of bacteria. Biochemical processes tend to be more stochastic when they involve molecules that occur in low numbers. As the number of molecules of a species gets closer to zero, small fluctuations have a significant impact on the total number of molecules present. RNAs are a good example of this, they degrade fast relative to other molecular species, such as proteins, so the number of RNAs tends to be low, making RNAs a major source of noise in gene expression [5]. If the number of all RNAs in a cell is increased, we would expect to see less variation in gene expression, which would lead to less variation in protein production.

The toehold switch sequence is made up of 3 Gs, a 12 nt linear "toehold" and then a 47 nt sequence that, when at equilibrium, forms a hairpin structure containing the 8 nt RBS inside a loop [55]. For the hairpin loop to occlude the ribosome, it needs to have a strong and stable secondary structure, so the hairpin has a low, negative free energy. The more negative the free energy of an RNA secondary structure, the harder it is to unfold because more energy is required to do so. For the sRNA to unfold the hairpin structure, the toehold region of the switch must be linear and unbound. When sRNA binds to the THS, initially, the first 12 nt of the sRNA binds easily to the linear toehold because it does not need to unfold the strong hairpin secondary structure. This initial binding to the toehold triggers a branch migration process, where a series of small intermolecular interactions unfolds the hairpin to reveal the RBS. Less energy is needed

to unravel the hairpin loop once the sRNA has already bound the toehold. Although the hairpin of the THS is stable, the toehold makes it energetically favourable for the sRNA to bind to it and trigger its unfolding.

We hypothesized that the RiboJ insulator upstream of the THS might interfere with the toehold, making it easier or harder for the sRNA to unfold the THS, which either increases or decreases the rate that the sRNA can bind. The speed that the hairpin folds also affects how well the THS represses translation. Based on simulations of bases pairing when RNA folds around an RBS to prevent translation, it was shown that RNA structures can have similar free energies and still fold and occlude ribosomes at different speeds, which affects the rate of protein production [43]. This effect can cause over 1000-fold variation in the rate of protein production. If the time taken for the switch to fold and prevent ribosomes binding the RBS is more than the time taken for a ribosome to bind to an RNA and initiate translation, then the probability of a ribosome successfully binding to an RBS and initiating translation is greater. If the RiboJ slows the rate of THS folding, the THS will not effectively stop ribosomes translating our gene of interest, so the rate of translation will be higher, even when no sRNAs are present [43]. Therefore, we propose that removing the RiboJ insulator would improve the TES's performance.

5.3 Increasing sRNA concentration

When the input and tuner promoters are most active, the concentration of the sRNA is expected to be less than half the concentration of the THS. Less than half of the THS molecules are bound by sRNAs and so less than half of the mRNAs can be translated. We wanted to increase the number of sRNAs that the TES can produce, so at high input and tuner promoter activities, there will be more sRNAs to bind to THSs which will enable more protein to be produced. We hoped this would increase the fold-change of the TES.

5.3.1 Modelling increased rate of sRNA transcription

To test our theory, we used the model derived in Chapter 4 to simulate how changing the input and tuner promoter activities and sRNA and mRNA association rates affected the fold change of the TES (**Figure 5.1**). All simulations were run with the model shown in **Eqs. (4.10 - 4.13)** with parameters shown in **Table 4.1**. Input promoter activity was simulated at rates that increased in logarithmic intervals from 0.0001 RNAP to 300 RNAP min^{-1}. Increasing the rate of sRNA transcription increased the TES's fold-change (**Figure 5.1A**). We see this trend at all rates of RNA association we tested. However, when the tuner promoter activity is low (<1 RNAP min^{-1}), increasing the tuner promoter activity doesn't affect the device's fold change (**Figure 5.1B**). When the tuner promoter activity increases above a certain point (>1 RNAP min^{-1}), the fold-change of the device increases significantly. This trend is seen at all rates of RNA association. As we increase the rate of RNA association, the increase in tuner promoter activity needed to

increase the fold-change of the device decreases. This shows that if we increase the rate of sRNA transcription enough, we should be able to increase the fold-change of the device.

Another problem with the performance of the THS was that a large percentage of individual cells from populations of cells grown at low and high input promoter activities showed similar fluorescence levels. The intersection between these two population distributions was more than 69% at all rates of sRNA transcription we tested (**Figure 4.9B**). We wanted to decrease the intersection between distributions by increasing the distance between them. We predicted that if we increase the rate of transcription of sRNA to increase the device's fold-change we will also decrease the intersection.

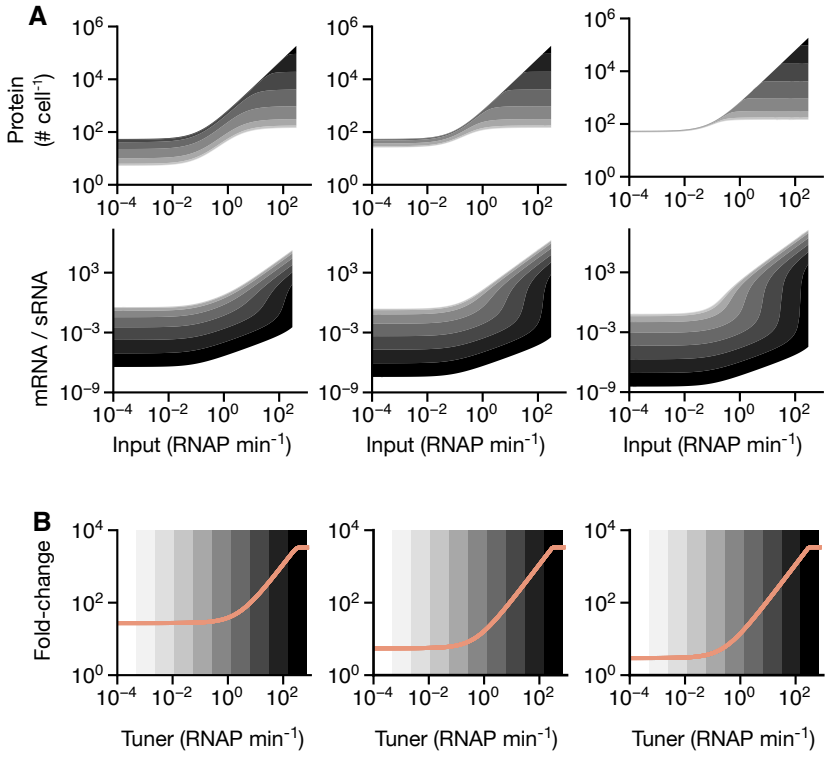


Figure 5.1: Simulations show how changing model parameters changes molecule concentrations. The rate of association between sRNA and mRNA was simulated at, from left to right panels, 0.0257, 0.257 and 2.57 complexes transcript^{-1} min^{-1}. (**A**) Top panel shows concentration of protein at steady state vs input promoter activity (input). Bottom panel shows changes in the ratio between mRNA and sRNA concentrations at steady state (mRNA/sRNA) vs. Input. (**B**) Changes in fold-change between low and high inputs with tuner promoter activity (tuner). Ranges of tuner promoter activity are shown in coloured bands where colour signifies intervals between: 0.0001, 0.0005, 0.0024, 0.012, 0.056, 0.27, 1.3, 6.4, 31, 150 and 730 RNAP $^{-1}$, from light to dark.

To test if increasing the rate of transcription of sRNA would decrease the intersection between two population distributions, we ran stochastic simulations of our model found in **Eqs. (4.10 - 4.13)** (see Chapter 3, Section 3.5). We ran simulations at both low and high rates of sRNA transcription and found that increasing the rate of transcription of sRNA increased the distance

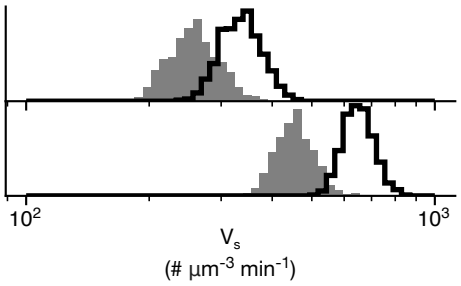


Figure 5.2: Population distributions from stochastic simulations of a tuneable expression system. 5000 stochastic simulations were run for the TES at low (1 RNAP min^{-1}) and high (1.5 RNAP min^{-1}) promoter activities, shown by grey and black outlined distributions respectively. This was repeated at low (1.5 RNAP min^{-1}) and high (5 RNAP min^{-1}) tuner promoter activities, shown on the top and bottom of the figure, respectively.

between the two population distributions as well as decreasing their variation (**Figure 5.2**). These effects reduced the intersection from 31% at low rates of sRNA transcription to 6% at high rate of sRNA transcription.

5.3.2 Design of a booster plasmid to increase sRNA transcription rate

To increase the rate that sRNA was transcribed we designed a sRNA booster plasmid. Cells were cotransformed with the TES and booster plasmids (**Figure 5.3**). The booster plasmid produces high levels of sRNA in response to IPTG. When cells harbouring the booster and TES are used together, the booster design ensures that more sRNA is produced than mRNA, ensuring maximal mRNA translation rates are achieved.

Two key features of the booster plasmid design ensure that high concentrations of sRNA are produced. First, we placed transcription of sRNA under the control of a strong P_{T7} promoter (**Figure 5.4**). P_{T7} promoters are transcribed by the viral T7 RNA polymerase (RNAP), which processes DNA 8–times faster than the native bacterial RNAP [89]. In the genome of our testing strain, *E. coli* BL21 Star (DE3), LacI represses transcription from a constitutive P_{LacUV5} promoter, preventing the transcription of the downstream gene coding for T7 RNAP. When IPTG is added to the system, it binds to LacI, changing its conformation in a way that prevents it repressing transcription and thus enabling the downstream T7 RNAP gene to be expressed and allowing it to transcribe the sRNA regulated by the P_{T7} promoter. Secondly, the booster plasmid was built with a pColE1 origin of replication, which has a copy number of 50-70 copies per cell [85] (**Figure 5.4**). This is 3–4.5 times higher than the ~15 copies per cell of the p15A origin used for the plasmid containing the TES device. Because we have more copies of the booster plasmid, we expect that the rate of transcription of sRNA will be higher than from the device plasmid (**Figure 5.3**).

5.3.3 Increased sRNA concentration improves system performance

We characterised the TES in cells cotransformed with the sRNA booster plasmid. Compared to the TES on its own, we saw an increase in protein production at higher rates of mRNA transcription (> 6.04 RPU). Unexpectedly, the rate of protein production at low rates of mRNA transcription (<0.5 RPU) decreased (**Figure 5.5A**). The most likely cause for this is that the sRNA booster expresses a second antibiotic selection gene, ampR, which give cells harbouring the booster resistance to the antibiotic ampicillin. The presence of a second antibiotic in the growth media, and the added burden of expressing an additional protein will yield reduced rates of protein production in the cells [9, 26, 27].

We calculated the fold-change for the device with the booster and found that at low tuner input (0.002RPU), fold-change increased from 14 to 227. This is a large increase, however, fold–change is still an order of magnitude lower than we expect based on how the sensor inputs to the device were characterised independently, where we saw a fold-change of over 1000 (Chapter 4, Section 4.4.1). Whilst fold–change increased at all tuner inputs, the biggest increase was seen at the lowest tuner input (a 16–fold increase at 0.003 RPU). Fold change increased by ~5–fold for tuner inputs between 0.03–0.9 RPU and ~2-fold at a tuner input of 2.61 RPU (**Figure5.5B**).

5.3.4 Effect of increased sRNA concentration on output distributions

When we compare population distributions of devices characterized with the booster plasmid against devices characterised without, the intersection decreases at all tuner inputs. The biggest improvement was at a tuner input of 0.9 RPU, when the intersection dropped by 77%. The smallest improvement was when the tuner input was 0.03 RPU and the intersection dropped by 25%. At all other tuner inputs, the intersection dropped by >40% (**Figure 5.5C**). The improvements are seen partly because the median output fluorescence at low inputs (0.002 RPU) decreases, reducing the intersection between distributions. For example, at low input and high tuner input promoter activities, the average median fluorescence in populations of cells with the booster was 1.7–times lower (**Figure 5.5A**). The intersection was further reduced due to the average

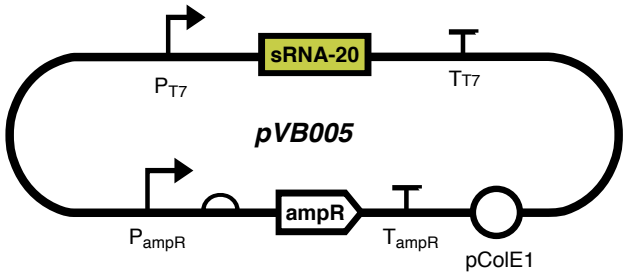


Figure 5.3: Design of sRNA booster. P_{T7} is the T7RNAP activated promoter and T_{T7} is its terminator. ampR is the gene encoding ampicillin resistance.

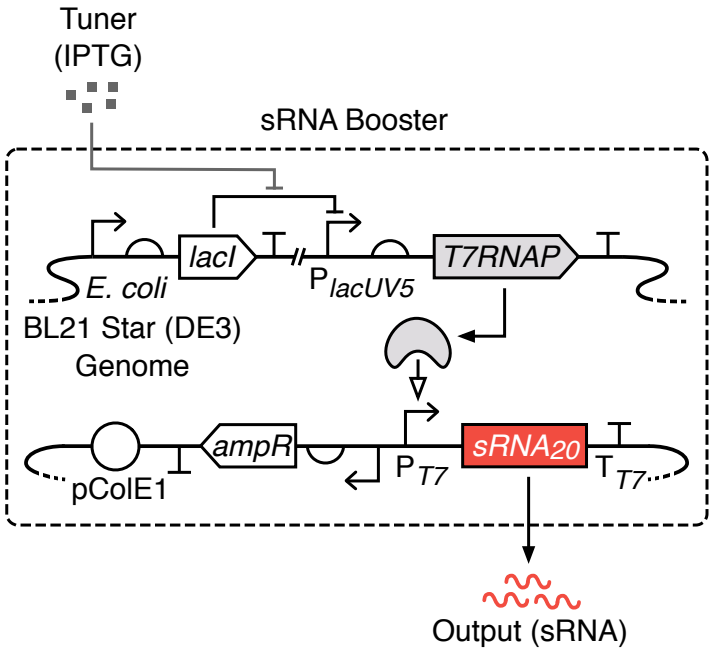


Figure 5.4: Design of sRNA booster T7 RNA polymerase (T7RNAP) is encoded on the hosts genomes, it transcribes sRNA from DNA found on a separate plasmid in response to IPTG.

median fluorescence at high input promoter activity (6.6 RPU) increasing. When input and tuner promoter activity were high (6.6 and 2.61 RPU, respectively) the median fluorescence was 1.6 times higher with the booster, which further separated distributions and reduced intersection. These changes can be seen from the response functions in **Figure 5.5A**, where the minimum output decreases and the maximum output increases for all tuner promoter activities.

Our model was useful for improving the design of the TES, however, if we compare modelling results to experimental results, we see that the model fails to capture key features of our experimental system. For example, our model suggests that as we increase the rate of sRNA transcription the fold-change of the device will remain the same or increase. Contrary to what our model suggests, in our experimental system the fold-change decreases as we increase the rate of sRNA transcription. The model fails to capture how tuning the rate of sRNA transcription varies protein production to different extents at low and high input promoter activities. This difference in behaviour suggests that we've failed to model key biological processes occurring in the TES or that the parameters we use are incorrect. However, as we suggest previously, if the THS is not working correctly in the TES, the TES would not behave as we expect based on model predictions.

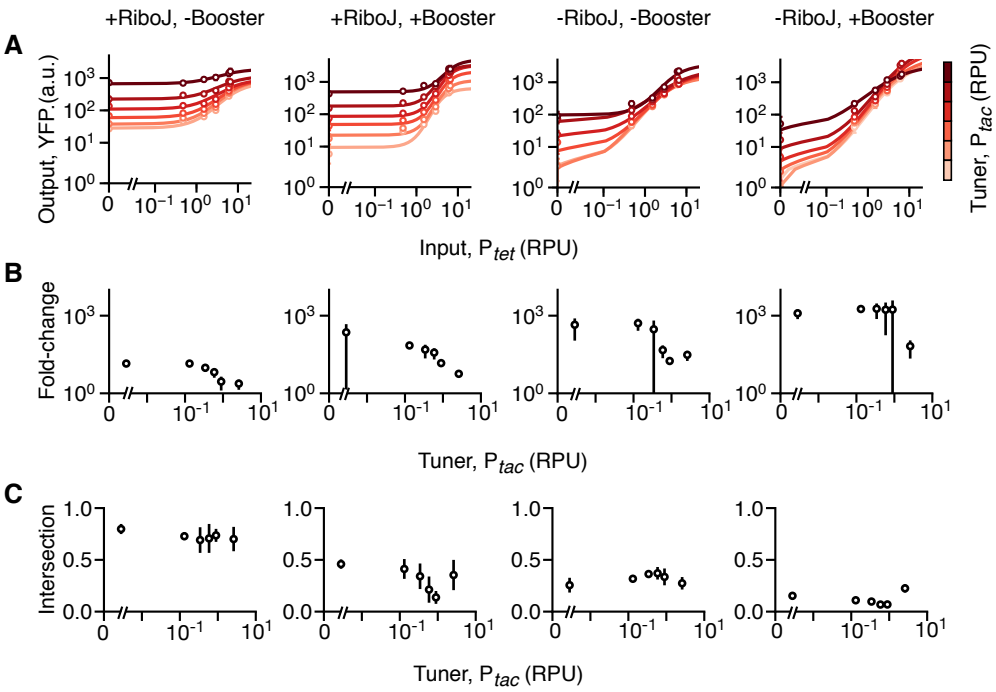


Figure 5.5: Comparison of tuneable expression system performance with and without sRNA booster plasmid. (**A**) Experimental response function for the TES. Lines show fitted hill functions for tuner promoter activities 0.003, 0.03, 0.15, 0.43, 0.9 and 2.61 RPU, colored from light to dark red. (**B**) Change in fold-change of device with change in tuner promoter activity. (**C**) Change in the fraction of intersection of fluorescence distributions between cells grown at high (6.6 RPU) and low (0.002 RPU) input promoter activities. For all plots, points are averages of three biological replicates, error bars show ± 1 standard deviation.

5.4 Removal of RiboJ insulator

5.4.1 Thermodynamic modelling

At low rates of sRNA transcription and high rates of mRNA transcription, the majority of THSs will not be sRNA bound and should remain folded around their RBSs, preventing translation of YFP. Therefore, at low rates of sRNA transcription we expect to see almost no protein production. However, experiments showed fairly high rates of protein production even when sRNA was produced at its lowest rates (**Figure 5.5**). This suggests that the THS was not folding correctly and so failed to block ribosomes. We suggested in Chapter 4 that the RiboJ insulator might interfere with toehold switch folding. Like the THS, the way in which RiboJ folds determines how it functions. The strong RNA secondary structure that the RiboJ forms could prevent the THS folding correctly, allowing YFP to be translated when it should have been repressed.

To test if the RiboJ secondary structure interferes with the THS, we modelled how pairing between bases in the THS causes it to fold at equilibrium [45]. We did this for the THS with and without cleaved upstream RiboJ insulators (for detailed method see: Chapter 3, Section 3.5.3).

We then modelled RNA folding with the sRNA present (**Figure 5.6**). We found that the THS structure was less stable with the RiboJ insulator upstream; the free energy of the THS was -40.5 kcal/mol with the RiboJ compared to -65 kcal/mol. The less stable the THS structure the easier it is for ribosomes to access the RBS. So, when the rate of sRNA transcription is low and the THS is folded, the rate of translation of genes regulated by a THS next to a RiboJ will be higher than for non-insulated genes. From our equilibrium base-pair model we see that the first 4 nts of the toehold section of the THS pairs to the RiboJ and are incorporated into a hairpin structure (**Figure 5.6**). The toehold makes it easier for sRNA to bind to and unfold the THS. As the first four nts of the toehold are incorporated into the RiboJ hairpin, more energy is needed for the sRNA to bind to the toehold, initiate the unfolding of the THS and activate translation. As a result, the sRNA will associate with the THS at a lower rate, leading to lower rates of translation.

The base-pair modelling approach we used predicts how bases will pair in RNA systems when they are in equilibrium, however they fail to explain how non-equilibrium processes, such as RNA folding, will affect our system. To address this, we modelled the time taken for a THS to fold and compared it to the time taken for a THS with an upstream cleaved RiboJ to fold (for method see: Chapter 3, Section 3.5.3). The kinetic model implements Monte Carlo procedures to simulate stochastic folding of RNA sequences [45]. We found that it took 7-times longer for a THS to fold into an equilibrium state when it had an upstream RiboJ. As the THS with the RiboJ

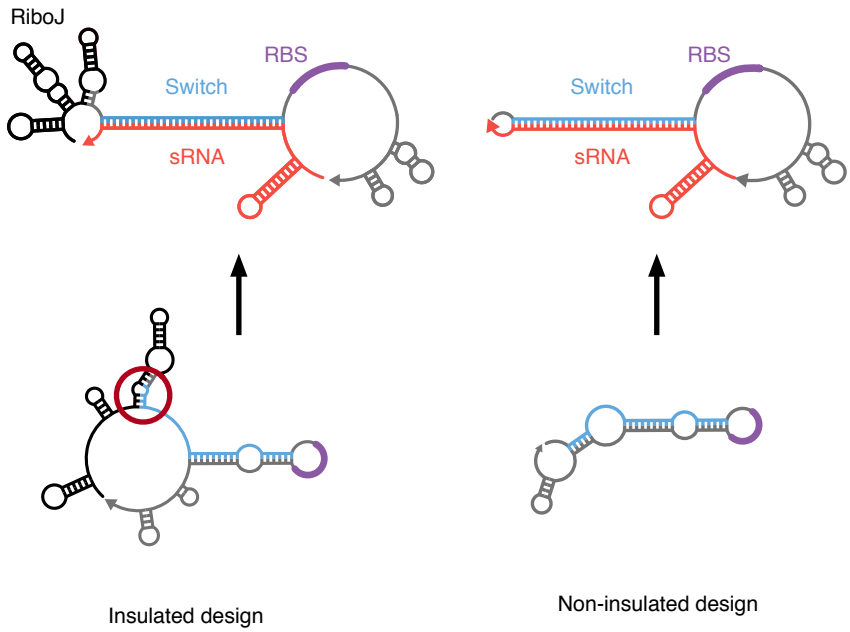


Figure 5.6: Predicted pairing of bases in the tuneable expression system at equilibrium. Secondary structure for insulated and non-insulated tuneable expression systems (TES) shown on the left and right, respectively. Bottom figures show closed, inactive toehold switches (THS), top figures show active THSs bonded with sRNAs. The red circle highlights the 4 nts of the toehold portion of the THS that are bonded to RiboJ.

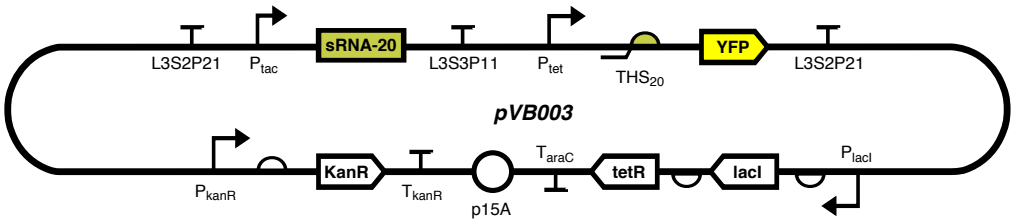


Figure 5.7: Tuneable expression system plasmid map without RiboJ insulator. Schematic diagram of the tuneable expression system plasmid without a RiboJ insulator (pVB003) in between P_{tet} and THS_{20}.

takes longer to fold, there is more time for ribosomes to initiate translation of downstream genes before the THS blocks translation, so the rate of protein production will be higher. When sRNA is not transcribed, we expect protein production to be repressed. These results could explain why the TES with the RiboJ still produces protein when the rate of sRNA transcription is at its lowest (0.002 RPU). Based on the findings from kinetic modelling and base-pair equilibrium modelling, if we remove the RiboJ from our original TES design, we expect to see less protein produced at low input promoter activities and more at high input and tuner promoter activities.

5.4.2 Removing RiboJ insulator improves performance

We wanted to see if the insulator was causing protein to be produced faster when the rate of sRNA transcription was low and slower when the rates of mRNA and sRNA transcription were high, so we removed the RiboJ insulator from the original TES design and tested the new variant. Compared to the original design, the output from the device when the input promoter activity was low (0.002 RPU), at all rates of sRNA transcription we tested, was lower (**Figure 5.8A**). This supports the hypothesis that RiboJ interferes with how the toehold switch performs by preventing it from sufficiently repressing translation of the associated gene.

As well as a decrease in protein production at low inputs, as we expected, we also saw an increase in protein production when the input was at its highest (6.6 RPU). This likely happens, as our model suggests, because the RiboJ binds to the toehold and slows down the rate of association between sRNA and toehold switch. By removing the RiboJ, we increased this rate of association, and in doing so, increase the rate of protein production (**Figure 5.8A**). These changes caused the fold change of the TES to increase across all rates of sRNA transcription tested (**Figure 5.8B**). Compared to the original device, the dynamic range more than doubled, the fold-change increased more than 10-fold at all rates of sRNA transcription and the fraction of intersection between YFP distributions decreased by more than 50% (**Figure 5.8**).

Next, we wanted to see if we could further improve the device's performance by using the sRNA booster in combination with the non-insulated TES variant. We cotransformed the non-insulated TES into cells with the booster plasmid and characterised their function. Compared to

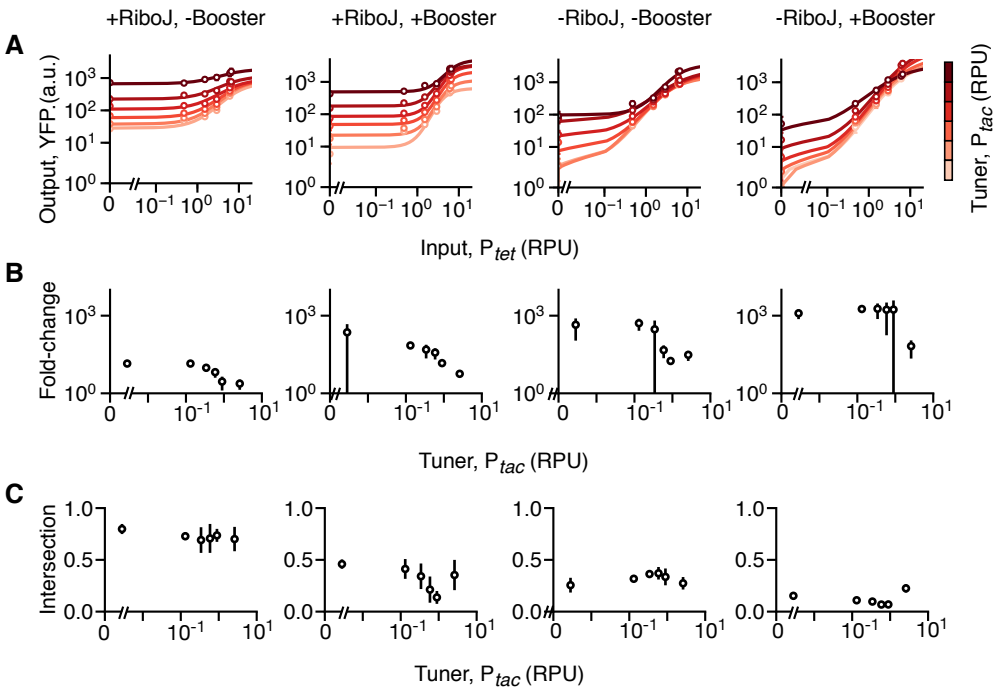


Figure 5.8: Comparison of tuneable expression system performances. (**A**) Lines show fitted hill functions for tuner promoter activities 0.002, 0.03, 0.15, 0.43, 0.9 and 2.61 RPU, colored from light to dark red. (**B**) Change in fold-change of device with change in tuner promoter activity (tuner). (**C**) Change in fraction of intersection of fluorescence population distributions between cells grown at high (6.6 RPU) and low (0.002 RPU) input promoter activities, with a change in tuner promoter activity. In all figures, points are averages of three biological replicates, error bars show ± 1 standard deviation.

the non-insulated design the dynamic range did not improve. At low input promoter activities it increased by 75%, but at high input promoter activities it decreased by 26%. In contrast, the fold-change increased more than 2-fold at all tuner promoter activities (**Figure 5.8B**).

As we saw when we added the booster plasmid to the insulated TES, the rate of protein production decreased at low input promoter activities across all tuner promoter activities. This decrease in protein production caused the increase in fold-change, as there was no significant increase in protein production at high inputs. At tuner inputs less than 0.03 RPU and at the lowest input promoter activity (0.002 RPU), protein production is almost the same as autofluorescence, showing that the THS tightly represses translation. However, at the lowest input promoter activity we still see a significant increase in protein production as we increase the tuner promoter activity.

At the lowest input promoter activity, we expected to see a small amount of mRNA transcribed due to the promoter leaking. If this was the case, an increase in tuner promoter activity would cause a small increase in the rate of protein production. What we see is a large increase in protein production as we increase tuner promoter activity from low to high at low input promoter activities. Increasing the tuner promoter activity from low to high (0.003–2.61 RPU) increases

Table 5.1: Summary of tuneable device performances

	Dynamic range[a,b](a.u.)		Fold-change[a,c](a.u.)		Intersection[a,d]	
Design	Low[e]	High[f]	Low[e]	High[f]	Low[e]	High[f]
Original	333 ± 53	877 ± 695	14 ± 1.7	2.4 ± 1.2	0.78 ± 0.06	0.69 ± 0.16
sRNA booster[g]	538 ± 51	2064 ± 1070	227 ± 297	5.7 ± 1.8	0.46 ± 0.04	0.35 ± 0.15
Non-insulated[h]	882 ± 134	2149 ± 409	445 ± 412	31 ± 16	0.26 ± 0.07	0.27 ± 0.06
Combined[i]	1550 ± 209	1712 ± 584	1236 ± 613	66 ± 54	0.15 ± 0.04	0.22 ± 0.04

a. Average of median values from flow cytometry are ± 1 standard deviation calculated from three biological replicates.
b. Dynamic range is the absolute difference in YFP fluorescence between low and high inputs (0.002 and 6.6 RPU).
c. Fold-change in YFP fluorescence (corrected for cell autofluorescence) for low and high inputs (0.002 and 6.6 RPU).
d. Fraction of intersection between flow cytometry YFP fluorescence distributions for low and high inputs (0.002 and 6.6 RPU) (3).
e. Performance measured for a low tuner input (expected $\mathbf{P}_{tac}$ activity = 0.002 RPU).
f. Performance measured for a low tuner input (expected $\mathbf{P}_{tac}$ activity = 2.61 RPU).
g. Original designs (**Figure 5.8**) with the sRNA booster system (**Figure 5.4**).
h. Design without RiboJ insulator.
i. Design without RiboJ insulator but with sRNA booster (**Figure 5.4**).

the output when the input promoter activity is lowest (0.002 RPU) by >28–fold for all of the TES variants. One explanation is that, when sRNA is transcribed the terminator after the sRNA sequence fails to stop RNAPs from transcribing. The RNAP reads through the terminator and transcribes the downstream THS and mRNA. As we increase the sRNA promoter activity, we increase the rate of sRNA transcription and also the rate of mRNA transcription. This would cause a significant increase in protein production at low mRNA promoter activities as we increase the sRNA promoter activity, as seen in our results (**Figure 5.8A**). Moreover, the intersection further decreased when we characterised the non-insulated TES with the booster (**Figure 5.8C**). A summary of all device performances can be seen in **Table 5.1**.

5.5 Discussion

In this chapter we used a deterministic model of the TES to explore how changing key design parameters affected its performance. This revealed that increasing the rate of sRNA transcription increased the fold-change of the TES. We then used stochastic modelling to show that increasing the rate of sRNA transcription decreases the fraction of intersection between distributions of output proteins in cells grown at low and high rates of mRNA transcription. To increase the rate of transcription of sRNA we built an sRNA booster plasmid which we used, alongside the TES

plasmid, to cotransform cells. As expected from our model, the TES with the booster produced proteins at a higher rate leading to a higher fold-change at all tuner promoter activities tested.

Next, we used thermodynamic models to simulate how the RiboJ insulator affects toehold switch function. Our model suggested that positioning a RiboJ RNA upstream of a toehold switch makes it less thermodynamically stable, preventing it from repressing translation as effectively. A kinetic model of RNA folding for the THS with and without an upstream RiboJ showed that the insulated THS takes 7-times longer to fold, meaning more ribosomes could translate a regulated gene before the THS represses translation. The thermodynamic models we used effectively predict RNA secondary structure for RNA strands *in-vitro*. However, due to the vastly different conditions RNAs are exposed to in a cell, model outputs aren't necessarily accurate at predicting RNA secondary structure *in-vivo* [39, 78]. Despite this, we expected removing the insulator would enable the THS to fold correctly and inhibit translation, so a non-insulated TES would produce protein at a lower rate at low input promoter activities.

We built a TES variant without a RiboJ and found it produced protein at a higher rate at high input promoter activities and a low rate at low input promoter activities, yielding a device with a higher fold-change. Whilst this supports our modelling results, another explanation for the high rate of protein production we were seeing at low input promoter activity, is that the abundance of RiboJ insulated mRNA is around two-fold higher than non-insulated mRNA [34]. It is thought that the terminal hairpin formed by the cleaved RiboJ, stabilizes mRNA, increasing its concentration and, therefore, the rate of protein production [22, 34, 83]. Whilst it is likely that the RiboJ interfered with TES function through the mechanisms our model highlighted, the RiboJ may have affected device performance in other ways. This has taught us that in general, if building a device with two or more neighbouring parts that rely on the secondary structure of their RNA to function, it is likely that the parts will interfere with each-other.

Finally, we tested the non-insulated TES with the booster plasmid. Compared with previous variants of the device, the new TES variant showed higher fold-changes and lower intersection for all tuner promoter activities tested (**Figure 5.8**). However, the rate of protein production still increased with tuner promoter activity. An explanation for this is that when the tuner promoter was active, the terminator after the sRNA sequence failed to terminate transcription. If this was the case, RNAP would have read through the terminator and transcribed the THS and mRNA in addition to the sRNA. Even when input promoter activity was at its lowest, mRNA would have been produced at higher rates as we increased the rate of sRNA transcription. This is what we saw from our experimental data. To verify if the increase in protein production happens due to transcriptional read-through, we could sequence the RNA in cells grown at the lowest input promoter activity and at varying tuner promoter activities [51]. If we see an increase in mRNA production as we increase tuner promoter activity this would confirm that we get transcriptional read-through. We could solve the problem of transcriptional read-through in one of two ways. First, we could remove the sRNA cassette from the TES plasmid and express it solely from the

booster plasmid. Second, we could rebuild the TES device plasmid, but with the sRNA cassette in the reverse direction.

Based on these experiences, we recommend using thermodynamic modelling when designing any device or circuits that use RNA based parts. These models are easy to use and have been shown to be accurate [45]. Using them at the design stage could save resources wasted on re-building and testing devices that fail due to RNA based parts behaving unexpectedly. In terms of the work we present here, if we wanted to design a new TES that incorporates a RiboJ insulator without affecting THS function, we could design a number of spacers that we could place between the RiboJ and THS. We could use our thermodynamic modelling approach to find the best spacer design before building and testing it.

The new TES variants performed significantly better than the original design. Their improved performance makes them more suitable for use in larger genetic circuits. Instead of going through design-build-test cycles, synthetic biologists could incorporate TESs into their original designs. This way, if a circuit breaks, we could use the TES to tune devices and find the optimal rates of gene expression for different genes in a circuit. Sensitivity analysis and optimization methods could be used to identify parameters for the system that would yield a functioning circuit [44, 103]. However, such an approach assumes that the TES is modular and that changing its inputs and output protein will not affect its behaviour. The TESs we characterised here use the same input promoters (P_{tet} and P_{tac}) and tuner (THS-20) to regulate the same protein (YFP). For the TES to be used in a larger circuit it will need to be used to regulate different proteins and will likely need different inputs. Removing the RiboJ insulator improved the TES's performance, but at the same time we lost the functionality of the insulator. If this optimized device were to be used with different input promoters it would alter the 5' untranslated region and likely alter the rate of translation of our gene of interest [83]. Furthermore, the THS we used contains a linker that adds 11 residues to the N-terminus of the regulated protein [55]. Whilst this does not appear to affect YFP function, it could hinder the function of proteins that are sensitive to changes in their N-termini.

To make it easier to connect devices in order to build larger circuits, device inputs and outputs are often standardized so that they use the same signal. For example, biodesign automation software like Cello uses transcriptional logic gates as modules. Each module's input and output uses RNAP flux at a promoter as a standard signal carrier [19, 100]. In the next chapter, we use the TES to build a tuneable NOT gate with standardized inputs and outputs. In the NOT gate, the TES regulates the production of a key protein in the gate (a repressor protein), allowing us to control a key feature of its function – the transition point between its on and off output states. NOT gates can be combined with other transcriptional logic gates to build larger circuits capable of performing a wider range of functions, so tuneable NOT gates are a first step towards building circuits that can be tuned and fixed using our TES.

CHAPTER 6

Applications of a Tuneable Expression System

6.1 Introduction

Some genetic devices rely on the expression of proteins such as transcription factors to implement basic logic functions. Many of these can be composed to carry out more complex decision-making tasks [100, 124]. One such commonly used device is a NOT gate, which has a single input and output [124]. Its function is to 'invert' the input such that the output is high if the input is low and low if the input is high. Such a behavior can be implemented using promoters as inputs and outputs, with the input promoter driving expression of a repressor protein that binds to a constitutive output promoter. When the input promoter is inactive, the repressor is not synthesized, so the output promoter is active. When the input promoter is active, the repressor protein is produced. This binds to the output promoter, blocking RNA polymerases (RNAP) from transcribing downstream genes, causing the output promoter activity to decrease when the input promoter activity increases.

When connecting gates to build a larger circuit, the output promoter of one gate acts as the input to the next. The relationship between a NOT gate's input and output is represented as a sigmoidal response function, where increases in input promoter activity decrease the output promoter activity. Connected logic gates propagate a Boolean signal, which can be either 'on' or 'off', through a circuit. For a signal to pass between gates, gates' response functions need to exhibit switching-behaviour; whereby, low and high input promoter activities, corresponding to 'off' and 'on' input states, when mapped onto a response function, illicit two distinctive output signals that are either above or below a threshold.

A problem faced when connecting two gates is that, if the transition point between on and off states of the first is too low, the output promoter activity may not span the transition point

(K-value), which leads to an output that is permanently on or off. In this case, the gates are said to be mismatched and so the input signal can't propagate through the circuit. One or more mismatched pairs of gates in a circuit can impact its function, causing it to break. Thus, for gates to be matched, the transition point of both gates needs to be at a 'sweet spot', such that signals can propagate robustly.

To get around the problem of mismatched gates, biodesign automation (BDA) software such as Cello matches response functions to make sure that the transition points of connected gates are compatible with each other [100]. Mismatched gates can be fixed by modifying regulatory elements in their design. In the case of repressor-based NOT gates, while the promoters cannot be easily modified, in bacteria the translation initiation rate can be varied by altering the ribosome binding site (RBS) for the repressor gene. Increasing the RBS strength causes more repressor protein to be produced for the same input promoter activity (number of transcripts), shifting the transition point to a lower value [100]. While such modifications can fix issues with device compatibility, they require reassembly of the entire genetic device. In this chapter we propose a different solution to the problem and use a tuneable expression system (TES) to vary the translation rate of a gene with a NOT gate, allowing us to vary its transition point from an on to off output states.

The way a genetic part, device or circuit behaves is dependent on the environment it is used in. Cello accounts for this by allowing a user to define a user constraint file (UCF), where the conditions that parts or devices were characterised in are stated. However, most genetic parts and devices are tested under standard lab conditions, at 37 °C in a standard, predefined media [100]. This is fine for designing a proof of concept circuit, however, for circuits to be able to work in real world applications, such as in industrial scale bioreactors, they need to be robust to changes in environmental conditions. Changing the environment cells are grown in changes the response functions of circuits they harbour. Whilst a circuit might work under standard lab conditions, it is likely to break when grown in the varied conditions cells are exposed to in industry.

For example, when cells are grown at industrial scales, they are subjected to a range of different glucose concentrations as they circulate around bioreactors. These differences in glucose can have significant effects on the behaviour of the microorganisms. If glucose concentration is too high, cells grow too fast and use up all of the available oxygen in the local environment, which in turn triggers a stress response that significantly affects their physiology and how they express genes [42]. Similarly, if there is insufficient glucose, cell growth rates will drop, affecting the rate of protein production in cells and therefore the overall process [74, 75]. These effects will likely change the way synthetic circuits behave.

This chapter is structured as follows. In Section 6.2 we present a model to predict the behaviour of a tuneable NOT gate. We use this model to inform the physical construction of a tuneable NOT gate. Then, in Section 6.3 we build and test four variants of our tuneable NOT gate design, corresponding to the four variants of the TES we tested in Chapter 5. In Section

6.4 we show that the concentration of glucose that cells harbouring the TES are grown in affects the rate that they produce proteins. Finally, we demonstrate that, for a fixed input, we can tune the rate of protein production in different glucose concentrations, to ensure the rate of protein production is constant across all concentrations of glucose.

6.2 Model guided design of a tuneable NOT Gate

We created a proof-of-concept tuneable NOT gate that integrates a TES to allow its response function, and crucially its transition point, to be altered after assembly. We chose an existing NOT gate design that uses the PhlF repressor to control the activity of the output P_{phlF} promoter [124]. Expression of PhlF was controlled by the TES (replacing the YFP reporter protein in the original TES design). Unlike the TES, the tuneable NOT gate uses promoters for both inputs and outputs allowing it to be easily connected to other devices that use RNAP flux as an input/output signal. By using a common signal carrier the tuneable gate can be incorporated into BDA software such as Cello [100].

The NOT gate uses the TES to tune the production of a repressor protein, so we used the output from the TES model to demonstrate how the concentration of repressor, P, changes with input and output promoter activities. Once translated, the repressor binds to a constitutive output promoter (Figure 6.1). As we increase the concentration of repressor, the activity of the output promoter, r_O, decreases. We captured this relationship using the Hill function:

$$r_O = \alpha_{min} + (\alpha_{max} - \alpha_{max})\frac{K^n}{K^n + P^n}\,, \tag{6.1}$$

where , α_{min} and α_{max} are the minimum (basal) and maximum promoter activities, respectively, K is the repressor concentration at which the output promoter activity is halfway between its minimum and maximum, n is the Hill coefficient (cooperativity), and P is the steady state concentration of the repressor protein. The full model for the TES mediated tuneable NOT gate is given by:

$$\frac{dM}{dt} = r_M - k_C^+ MS + k_C^- C - \gamma_M M \tag{6.2}$$

$$\frac{dS}{dt} = r_S - k_C^+ MS + k_C^- C - \gamma_S S \tag{6.3}$$

$$\frac{dC}{dt} = k_C^+ MS - k_C^- C - \gamma_C \tag{6.4}$$

$$\frac{dP}{dt} = \beta_M M + \beta_C C - \gamma_P \tag{6.5}$$

$$r_O = \alpha_{min} + (\alpha_{max} - \alpha_{max})\frac{K^n}{K^n + P^n} \tag{6.6}$$

where M, S, C and P are the concentrations of mRNA, sRNA, complex (formed from mRNA and sRNA binding) and protein, respectively. M and S are transcribed at rates of r_M and r_S,

Table 6.1: Parameters used to model a tuneable NOT gate.

Name	Description	Value(s)	Unit	Ref.
α_{min}	Minimum (basal) output promoter activity	0.0015	RNAP min^{-1}	[72]
α_{max}	Maximum (basal) output promoter activity	500	RNAP min^{-1}	[72]
K	Repressor concentration where output promoter activity of the NOT gate is halfway between its minimum and maximum	4.76	proteins cell^{-1}	[85]
n	Hill coefficient for PhlF repressor	4	-	[85]

respectively. k_C^+ and k_C^- are the reaction rate constants for complex formation and dissociation, respectively. β_C is the reaction rate constant for translation of mRNA. M, S, C and P are degraded at a rate proportional to their concentration and rate constants, γ_M, γ_S, γ_C and γ_P, respectively. P is solving **Eqs. (6.2 - 6.5)** at steady state.

To model the NOT gate, the outputs from the simulations we ran for the TES (Chapter 4), were used to calculate the expected repressor concentrations to be used in **Eq. (6.6)**. We wanted to better understand how a NOT gate built with the repressor protein PhlF would behave so we chose biologically realistic parameters from the literature to assess how PhlF would repress the activity of it's related output promoter (**Table 6.1**).

Using this model, we simulated how steady state concentration of repressors and output promoter activity change as we varied input and tuner promoter activities, and association constants for the hybridization of sRNA and mRNA. We found that, as we increased input promoter activity there was a non-linear decrease in output promoter activity. We saw this at all rates of mRNA and sRNA association we tested (**Figure 6.2A**).

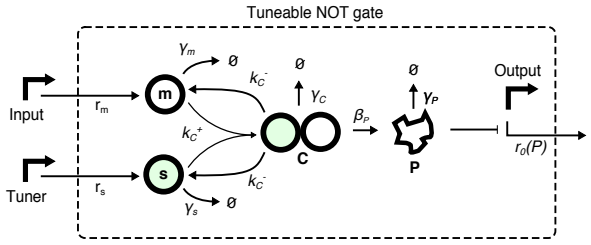


Figure 6.1: Key processes in a tuneable NOT gate. Parameters r_m, r_s and $r_0(P)$ are the input promoter activity, tuner promoter activity and output promoter activities, respectively. k_C^+, k_C^- and β_P are the rate constants for association of mRNA and sRNA, dissociation of mRNA and sRNA and translation of repressor protein, respectively. γ denotes a degradation constant, where the subscript represents the biochemical species being degraded. m, s, C and P are the mRNA, sRNA, complex formed between sRNA and mRNA and repressor protein, respectively.

There was a non-linear decrease in the fold-change between output promoter activities at low and high input promoter activities as we increased the tuner promoter activity. As tuner promoter activity was increased from 0.0001 to 730 RNAP min^{-1}, fold-change decreased: 10,000, 50 and 2–fold, at rate constants of sRNA and mRNA association of 0.0257, 0.257 and 2.57 complexes $transcript^{-1}$ min^{-1}, respectively (**Figure 6.2C**).

Increasing the rate of association reduced the range of output promoter activities over which we can tune the NOT gate. At low input promoter activities (0.0001 RNAP min^{-1}), increasing tuner promoter activity from 0.0001 to 730 RNAP min^{-1} caused the output promoter activity to decrease 6800, 20 and 1.7-fold at mRNA-sRNA association rates of 0.0257, 0.257 and 2.57 complexes $transcript^{-1}$ min^{-1}, respectively. As the rate of association between sRNA and mRNA increases, mRNA binds sRNAs at a faster rate, so for any input or tuner promoter activity, there are more mRNA molecules bound by sRNA that can be translated, leading to higher repressor concentrations.

6.3 In-vivo performance of a tuneable NOT gate

Based on these findings we built a tuneable NOT gate that uses a TES to regulate the production of PhlF (**Figure 6.3**). We constructed two variants. The first had a RiboJ insulator upstream of the THS; we refer to this variant as the insulated NOT gate. The second is a non-insulated variant, where the RiboJ insulator upstream of the THS has been removed. To measure the activity of the output promoter P_{phlF}, we placed a YFP reporter cassette downstream of the promoter (**Figure 6.3**). The cassette uses a RiboJ insulator upstream of the YFP to insulate translation of the YFP, so its rate is less affected by the 5'-untranslated region upstream of the YFP transcript.

Using the YFP cassette allowed us to convert NOT gate output fluorescence into relative promoter units (RPU) by comparing output fluorescence to the fluorescence of a standard plasmid (pAN1717) which uses a constitutive promoter to drive expression of the same YFP cassette [100]. Expressing promoter activities in RPUs, allowed us to compare them to other measurements found in the literature (Chapter 3, Materials and Methods) [19, 100].

We tested the proposed design *in-vivo* by transforming *E. coli* cells with the insulated (pVB002) and non-insulated variants (pVB004) of the tuneable NOT gate and performed additional co-transformations for each variant with the sRNA booster plasmid (Chapter 5). As with the TES, cells harboring the devices were grown in varying concentrations of aTc and IPTG, to modify the activity of the input and tuner promoters, respectively. The activity of the output promoter was inferred by measuring the fluorescence of single cells using flow cytometry, then converting this data into RPU units (Chapter 3, Section 3.4.5)

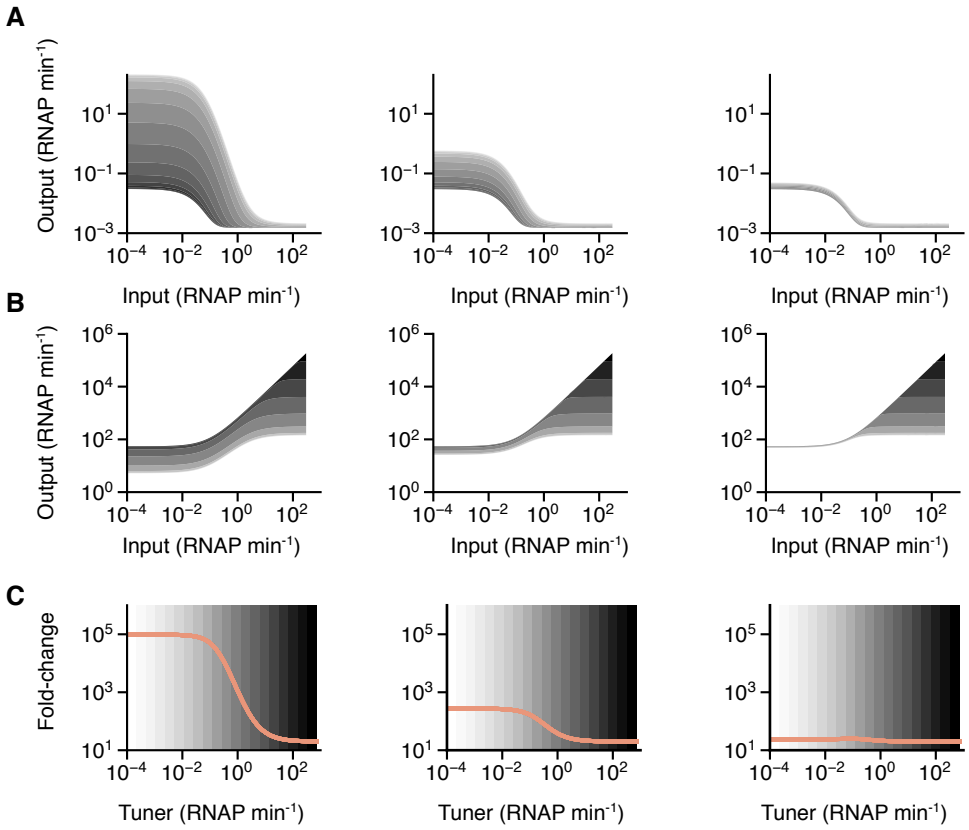


Figure 6.2: Deterministic modelling of a tuneable NOT gate. The rate of transcription of mRNA (Input) was increased from 0.0001 to 300 RNAP min^{-1}. The rate of sRNA transcription (Tuner) was increased in intervals between between 0.0001, 0.0005, 0.0024, 0.012, 0.056, 0.27, 1.3, 6.4, 31, 150 an 730 RNAP $^{-1}$, with changes in interval denoted by changes in band color from light to dark. The rate of association between sRNA and mRNA were modelled at, from left to right, 0.0257, 0.257 and 2.57 complexes $transcript^{-1}$ min^{-1} (**A**) Change in output promoter activity with input. (**B**) Concentration of repressor proteins (repressors) against input. (**C**) Fold change between NOT gate output promoter activity when input promoter activity is 0.0001 and 300 RNAP $^{-1}$, plotted against tuner activity.

6.3.1 Modifying repressor production tunes NOT gate transition

We plotted the relationship between output promoter activity and input promoter activity for varying levels of tuner activity (**Figure 6.4A**). Similar to our modelling results, the experimental response functions showed a negative sigmoidal shape. We fitted the data to a Hill function to find the transition points, K-value, of each response function. For the insulated gate without the sRNA booster plasmid, we found that there was a 7-fold difference in transition points between devices used at the highest and lowest tuner promoter activities (**Figure 6.5**). Increasing tuner promoter activity led to transitions at a lower input. The range of transition points achieved by our insulated gate covered a high proportion (35%) of the largest collection of repressor-based NOT gates built to date (total of 20 variants) (**Figure 6.5**) [100].

These results demonstrate that the TES could be used to construct a NOT gate where

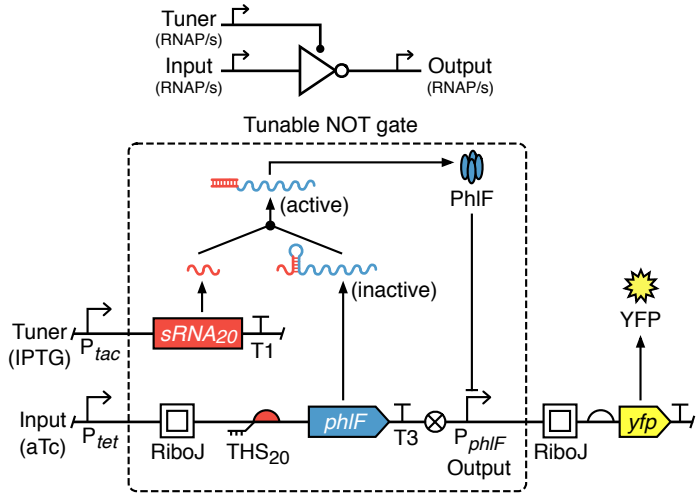


Figure 6.3: Tuneable genetic NOT gate implementation. Schematic of the tuneable NOT gate (top) and genetic design of the gate (bottom). The output promoter P_{phlF} drives production of a yellow fluorescent protein (YFP). T1 and T3 are transcriptional terminators L3S3P11 and ECK120033737, respectively.

transition point can be varied to improve compatibility with other genetic devices. However, tuning the gate came at a cost. As we predicted from our model (**Figures 6.2C**), increasing the tuner promoter activity decreased the fold-change between output promoter activities associated with low and high input promoter activities (**Figures 6.4A and B**). In our experimental system, increasing tuner promoter activity from 0.002 RPU to 0.9 RPU resulted in a 78% decrease in fold-change. As we increased tuner promoter activity there was a sigmoidal increase in the intersection of overlap between output YFP fluorescence distributions for cells grown at high (1.51 RPU) and low (0.0001 RPU) input promoter activities (**Figures 6.4A and C**). This increase was 4.5-fold as tuner promoter activity increased from 0.002 RPU to 0.9 RPU.

Once the tuner promoter activity reached a maximum working limit, increasing it any further made the fold-change too low and intersection between on and off states too high for the gate to be useful. Once this limit was exceeded, it became impossible to tell if in cells in a population are on or off.

6.3.2 Insulator and sRNA booster plasmid affect gate performance

We attempted to improve the performance of the tuneable NOT gate by characterising the insulated gate variant in cells co-transformed with the booster. The sRNA booster considerably improved the performance of the TES, so we thought it might improve the performance of the NOT gate too. However, compared with the original gate, the fold-change decreased at all tuner promoter activities we tested (**Figure 6.4B**). At low (0.002 RPU) and high (2.61 RPU) tuner promoter activities, fold-change decreased by 3% and 38%, respectively (**Table 6.2**). The range of transition points also decreased from 7-fold to 5-fold (**Figure 6.5**). The inclusion of the sRNA booster likely increased overall PhlF concentrations as the transition points shifted to far below

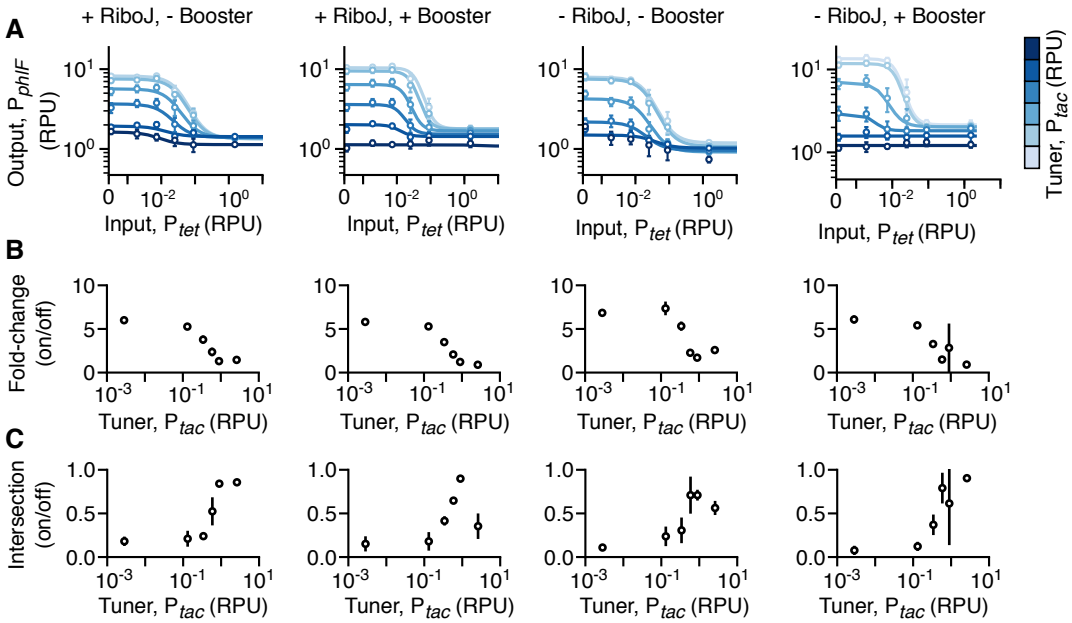


Figure 6.4: Characterisation of tuneable NOT gates. Points are averages of three biological replicates and error bars show ±1 standard deviation. (**A**) Steady state response functions of tuneable NOT gate variants. Lines are fitted hill functions at a single tuner promoter activity (light–dark: 0.002, 0.03, 0.15, 0.43, 0.90, 2.61 RPU). (**B**) Change in fold change between output promoter activities when input promoter activity is low (0.0001 RPU) and high (1.5 RPU), with tuner promoter activity. (**C**) Change in the intersection between output fluorescence population distributions with tuner promoter activity at low and high input promoter activities.

what had been seen for the original design. This would make this specific design of value for uses where a weak input signal needs to be inverted and amplified simultaneously.

We further improved the performance of the TES by removing the RiboJ insulator upstream of the THS part, which interfered with THS performance by preventing it from folding correctly and occluding ribosomes. The misfolded THS was also unable to be fully activated by sRNA. Removing the insulator decreased the rate of protein production at low input promoter activities and increased the rate of protein production at high input promoter activities, which worsened the TES's performance.

Based on how removing the RiboJ insulator improved TES performance, we thought the non-insulated NOT gate would perform better than the insulated variant. We expected the non-insulated TES embedded in the NOT gate to produce less PhlF at low inputs and more PhlF at high inputs, resulting in a higher output promoter activity at low inputs and lower output promoter activities at high inputs. This would in turn increase the device's fold-change compared to the insulated NOT gate.

We tested the non-insulated variant of the tuneable NOT gate, both without and with the booster. Compared to the insulated variant, the fold-change of the non-insulated variant was 13%

higher at a low tuner promoter activity (0.002 RPU) and 17% higher at a high tuner promoter activity (2.61 RPU) (**Table 6.2**, **Figure 6.4B**). The intersection at low and high tuner promoter activities decreased by 42% and 15%, respectively (**Table 6.2**, **Figure 6.4C**). The fold-change and intersection for the non-insulated design increased and decreased, respectively, because the output promoter activity at high input promoter activity was lower than in the insulated device. When comparing the non-insulated design to the insulated design at high input promoter activity (when we expect low outputs), the output promoter activity was 9.5% and 11% lower at low and high tuner promoter activities, respectively.

The output promoter activity from the non-insulated variant with the booster was 60% higher at low input and tuner promoter activities than from the insulated variant (**Figure 6.4A**). Compared to the insulated variant with the booster, the non-insulated TES with the booster produced 20% more protein at low input and tuner promoter activities (**Figure 5.5**). Despite the increase in output promoter activity at low input and tuner promoter activities, the fold-change of the non-insulated NOT gate decreases when it is expressed with the booster. At low and high tuner promoter activities, fold-change is 10% and 24% lower, respectively. These decreases likely occur because at high input promoter activity, when output promoter activity should be low, the output promoter activity is higher in the non-insulated gate with the booster for almost all tuner promoter activities we tested (**Figure 6.4**).

We chose to use the PhlF repressor in our NOT gate because in previous works, PhlF based NOT gates have been shown to perform well, exhibiting large dynamic ranges and fold-changes [100, 124]. For example, Nielsen et al., built three variants of a PhlF NOT gate with each using

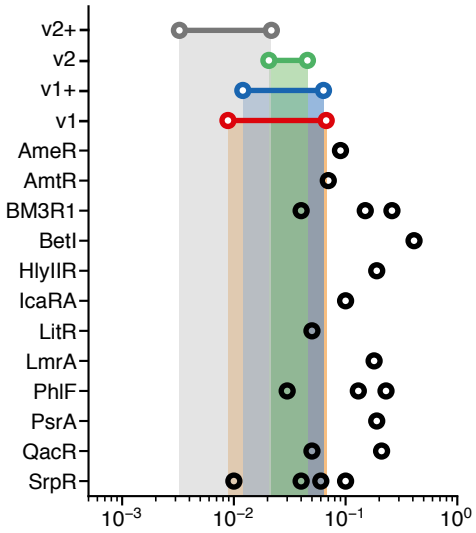


Figure 6.5: Comparison of NOT gate transition points. The range of transition points (K-values) of the insulated gate (v1), insulated gate with booster (v1+), non-insulated gate (v2) and non-insulated gate with booster (v2+) are compared with transition points of repressor based NOT gates used in Cello [100].

Table 6.2: Performance summary of the NOT gate designs

	Dynamic range[a,b](a.u.)		Fold-change[a,c](a.u.)		Intersection[a,d]		
Design	Low[e]	High[f]	Low[e]	High[f]	Low[e]	High[f]	K range
Original	8.15 ± 0.6	1.66 ± 0.134	6.0 ± 0.147	1.45 ± 0.09	0.19 ± 0.046	0.84 ± 0.02	0.009–0.067
sRNA booster[h]	10.4 ± 0.75	1.02 ± 0.31	5.8 ± 0.31	0.9 ± 0.25	0.13 ± 0.07	0.85 ± 0.02	0.012–0.063
Non-insulated[i]	8.24 ± 0.91	1.76 ± 0.39	6.8 ± 0.28	1.7 ± 0.23	0.11 ± 0.03	0.71 ± 0.08	0.021–0.046
Combined[j]	13.09 ± 1.46	1.12 ± 0.08	6.09 ± 0.56	0.91 ± 0.06	0.08 ± 0.05	0.9 ± 0.03	0.003–0.022

a. Average values are shown ± 1 standard deviation calculated from flow cytometry data for three biological replicates.
b. Dynamic range calculated as the absolute difference in YFP fluorescence between low (0.002 RPU) and high (1.5 RPU) inputs).
c. Fold-change in YFP fluorescence (corrected for cell autofluorescence) for low (0.002 RPU) and high (1.5 RPU) inputs.
d. Fraction of intersection between the flow cytometry YFP fluorescence distributions for low (0.002 RPU) and high (1.5 RPU) inputs (3).
e. Performance measured for a low tuner input (0.002 RPU).
f. Performance measured for a high tuner input (2.61 RPU).
g. Range of K values from Hill functions fitted to experimental data.
h. Original designs (**Figure 5.8**) with the sRNA booster system (**Figure 5.4**).
i. Design without RiboJ insulator.
j. Design without RiboJ insulator but with sRNA booster (**Figure 5.4**).

a different RBS [100]. For the three variants, the maximum activity of the output promoter P_{phlF} was between 3.9 and 6.8 RPU and minimum activity was between 0.01 and 0.02 RPU. The fold-changes of the devices ranged between 205 and 390, two orders of magnitude higher than we see from the tuneable NOT gates. The maximum output promoter activities from our tuneable NOT gate ranged from 8 to 13 RPU, values higher than those seen in Nielsen et al.'s gates. The output promoter activity at high inputs, for all our tuneable NOT gate designs was always at least 37-times higher than for their gates, suggesting that the tuneable NOT gates do not repress the P_{phlF} promoter activity as well.

6.4 Effect of glucose concentration on device performance

Even if a circuit is well designed to combine gates that match and then tested and verified in the lab, changing the environment a circuit is used in could change the shape of the circuit's response function, causing devices to mismatch and circuits to break. Despite this, we still characterise genetic parts in stable environmental conditions. For genetic circuits to be used in the real world, they need to be able to adapt to a range of environmental conditions. The first step towards building circuits that are robust to changes in environmental conditions is to build tuneable parts and devices that can adapt or be tuned depending on the environment they're in. To validate this approach, we need to start testing circuits in a range of different conditions that represent their end-use environments. In this way, we can discover how well circuits perform in different

environments. We can use this new understanding to build better genetic parts, devices and circuits.

The environment that cells are grown in directly affects their physiology and how they express genes. In this section, we investigated how an important variable in industrial bioreactors, glucose concentration, affects the performance of the TES. We did this to show how glucose concentration that cells are grown in affects the performance of our device and to demonstrate that the rate of gene expression from the TES can be tuned so that its behaviour remains constant across conditions. Because rates of protein production and device performance are directly related to growth rates, we initially investigated the effect of glucose concentration on the growth rate of cells harbouring the TES.

6.4.1 Cells grow faster in high glucose concentrations

To find out how glucose concentration affects the growth of cells, we measured the optical density of cells harboring the best performing TES design, the non-insulated variant, over time as they grew in glucose concentrations that varied over a 1000-fold range. We fitted the growth curves to a mathematical model and calculated the maximum growth rate of cells grown at each concentration of glucose by numerically calculating the maximum gradient of the fitted curves (Chapter 3, Section 3.4.8) (**Figure 6.6B**) [35] .

Experiments showed that the growth rate of cells increased more than 2-fold when the

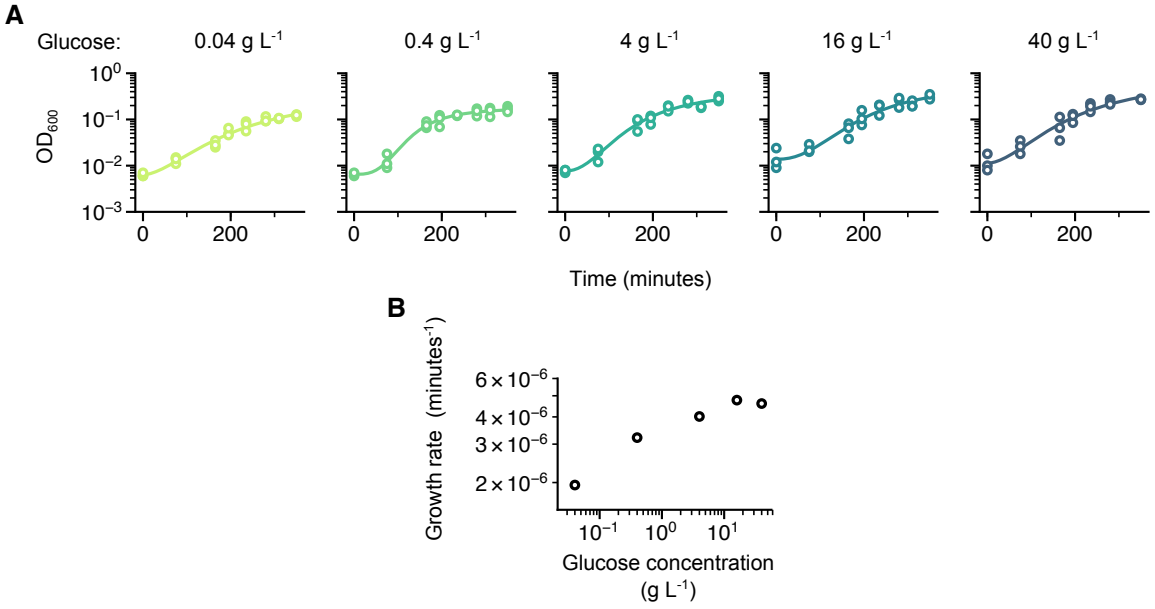


Figure 6.6: Effect of glucose concentration on cell growth. (**A**) Growth curves for cells grown over 350 minutes. Glucose concentration that cells were grown in is marked above each plot. Coloured circles are technical replicates. Coloured lines show a fitted mathematical model [35]. (**B**) Maximum growth rate, is plotted against glucose concentration.

concentration of glucose they were characterised in was increased between 0.04–4 g L^{-1}. However, increasing glucose concentration above 4 g L^{-1} had little effect on growth rate. Increasing glucose concentration to 16 and 40 g L^{-1} only resulted in growth rates 2.45 and 2.36–fold higher than at 0.04 g L^{-1}. These differences in growth rates show that, when glucose concentration is between 0.04 and 4 g L^{-1}, cells grow faster at higher concentrations of glucose (**Figure 6.6B**). Above 4 g L^{-1}, the effect of increasing glucose concentration on growth rate appears to plateau, and increasing glucose concentration has little effect. This indicates that glucose is a growth limiting substrate until glucose concentration is greater than 4 g L^{-1}.

6.4.2 Glucose concentration affects the performance of the tuneable expression system

We ran further experiments to see if cells produced protein at different rates depending on the concentration of glucose they were grown in. To do this, cells containing a TES (the non-insulated variant) were grown in different concentrations of aTc and IPTG, to vary the input and tuner promoter activities, respectively. Three different concentrations of glucose (0.04, 0.4 and 4 g L^{-1}), were used. Cells were sampled at 4 time points over a 3 hour period to see how the rate of YFP production changed over time using flow cytometry. The rate of protein production was given by the equation:

$$\frac{dP}{dt} = \beta - \gamma P \,, \tag{6.7}$$

where β is the rate of protein production, γ is the rate of protein loss and P is the concentration of protein. At steady state the equation becomes:

$$\beta = \gamma P \,. \tag{6.8}$$

We can calculate the rate of protein production by multiplying the concentration of protein by the rate of protein degradation and dilution. Because proteins are stable at the time scales considered, protein is mainly lost through dilution due to cell division [99]. When cells divide, all the proteins in a single cells are split, halving the number of proteins in each cell. We calculated the relative rate of protein production between samples by multiplying YFP fluorescence by growth rate. At all input and tuner promoter activities and time points we tested, the rate of protein production was lowest for cells grown in media containing the lowest glucose concentration, 0.04 g L^{-1} (**Figure 6.7**). We looked at protein production when cells were growing in exponential phase (at 120 minutes), as at this point, the rate of change in protein production is at steady state. We see that compared to cells grown in media with 0.04 g L^{-1} of glucose, the rate of protein production increased by between 25–174% and 12–140% for cells grown in 0.4 and 4 L^{-1} of glucose, respectively (**Figure 6.7B**).

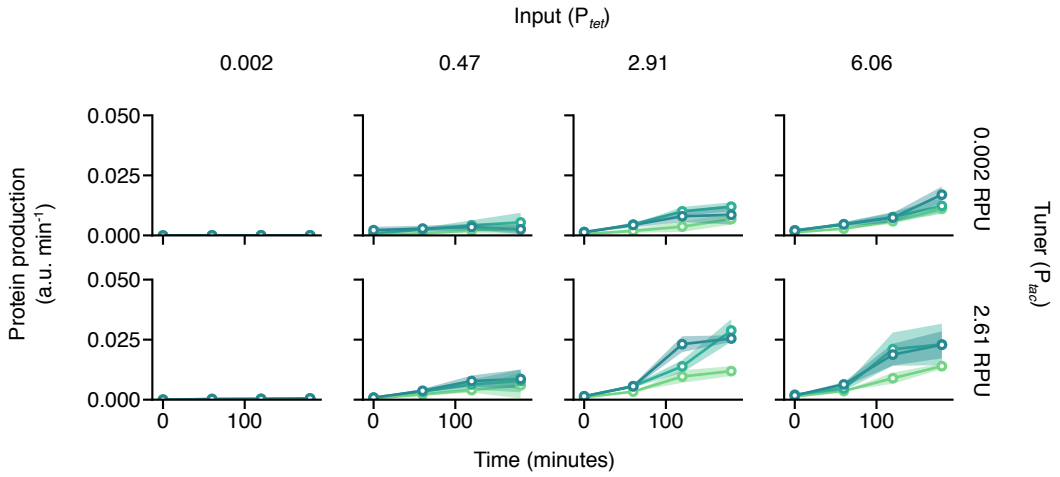


Figure 6.7: Effect of glucose concentration on protein production. (**A**) Change in protein production rate in cells grown in glucose concentrations of 0.04 (light-green), 0.4 (dark-green) and 4 g L^{-1} (blue), over 180 minutes. (**B**) Percentage difference in protein production rate compared to cells grown in 0.04 g L^{-1} of glucose. Circles are averages of three biological replicates, error bars show ± 1 standard deviation.

Growing cells in different concentrations of glucose has no apparent effect on the steady state response function of the TES (**Figure 6.8**). However, there are several difference in the performance. At low tuner promoter activities (0.002 RPU), the TES characterised in glucose concentrations of 0.4 and 4 g L^{-1} had fold changes between the rates of protein production at low and high inputs 40 and 36 times higher compared to when they were characterised in 0.04 g L^{-1} (**Figure 6.8B**). At high tuner promoter activities (2.61 RPU), the fold change was only 1.43 and 1.05 times higher when characterised in 0.4 and 4 4 g L^{-1}, compared to at 0.04 g L^{-1} (**Figure 6.8B**). Thus, the concentration of glucose clearly affects the performance of the TES, with the biggest effect seen at low tuner promoter activities, when the TES produces sRNA at a lower rate.

If a large circuit was used in conditions where glucose concentration varied, the device's performance could vary. A potential solution to this problem is to use a TES. The TES allows us to tune the rate of protein production, so that the device produces protein at the same rate in any glucose concentration. For example, at an input promoter activity of 6.06 RPUs, the TES can be used to tune the rate of protein production to be 0.009 a.u. min^{-1}, irrespective of how changing concentrations of glucose effect protein production. To do this we would adjust the tuner promoter activity by varying the relevant inducer to dynamically alter the rate of protein production to be 100%, 10% or 14% of the maximum rate that protein is produced at in cells grown in 0.04, 0.4 and 4 g L^{-1} of glucose, respectively (**Figure 6.8C**).

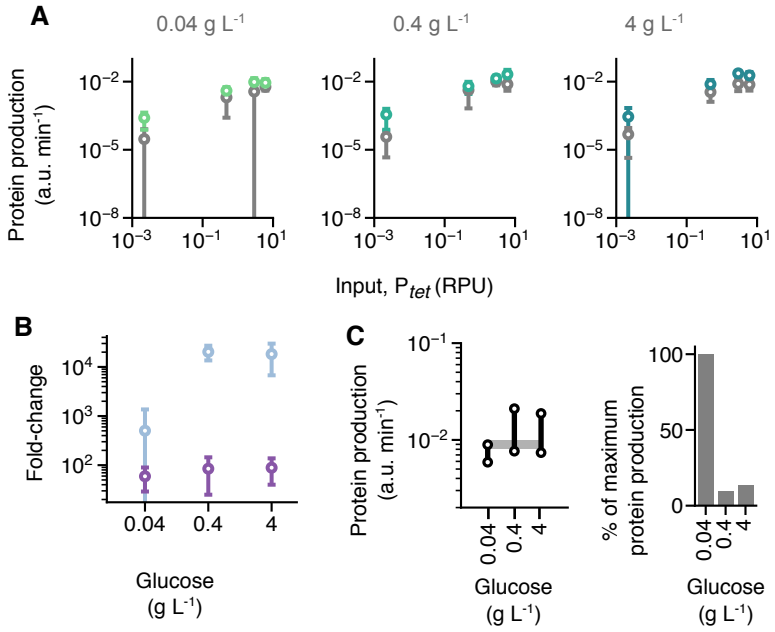


Figure 6.8: Effect of glucose concentration on device performance. Performance of TES grown in varying concentrations of glucose 120 minutes after induction. Circles are averages of three biological replicates, error bars are ± 1 standard deviation.(**A**) Steady state response function of tuneable expression system (TES) grown in varying glucose concentrations that are shown above each plot. Grey markers and coloured markers show protein production at low (0.002 RPU) and high (2.61 RPU) tuner promoter activities, respectively. (**B**) Change in fold-change in rate of protein production for TESs grown at high (6.06 RPU) and low (0.002 RPU) input promoter activities with glucose concentration. Light and dark purple denote low (0.002 RPU) and high (2.61 RPU) tuner promoter activities, respectively. (**C**) Range of protein production rates over which the TES can be tuned at different concentrations of glucose (left), the percentage of the maximum rate of protein production that the TES needs to be tuned to produce protein at a rate of 0.009 a.u. min^{-1}, in different concentrations of glucose (right).

6.5 Discussion

In this chapter, we built and tested four variants of a tuneable NOT gate: an insulated design, where the RiboJ insulator was placed upstream of the THS that regulates translation; a non-insulated design, where the RiboJ upstream of the THS had been removed and both of the first two variants were also co-transformed and tested with a sRNA booster plasmid. We showed that it is possible to tune the transition points of the tuneable NOT gates (**Figure 6.5**). The transition point for the insulated device without the booster could be tuned over the largest range, with the highest transition point being 7-fold higher than the lowest, a range of transition points that spans 35% of the largest library (made up of 20 variants) of repressor based NOT gates built to date [100]. To decrease the transition point, we had to increase the tuner promoter activity. However, for all four variants, as we increased the tuner promoter activity the fold-change of the gate decreased significantly. Comparing the output promoter activity of our NOT gate to data found in the literature, we found that the fold-change is low because the output promoter

activity is too high at high input promoter activities. This could be caused by RNAPs recruited to transcribe PhlF continuing past the terminator and transcribing the downstream YFP. However, because PhlF binds to P_{phlF} downstream of the terminator, the repressors might physically block RNAPs that have read through an upstream terminator, preventing transcriptional read-through.

PhlF represses transcription by binding upstream of a transcriptional start site which prevents RNAP from binding to DNA and initiating transcription [124]. Whilst the repressor prevents RNAP forming a complex with the DNA, it is not clear whether the repressor will inhibit an RNAP which has initiated transcription upstream, from elongating a transcript. However, reports in the literature suggest that repressors do not halt transcription of genes after transcription has already been initiated [106, 119].

We only see a low output promoter activity when the rate of PhlF production is at its highest (when input and tuner promoter activities are highest), so the terminator that halts transcription of PhlF prevents transcriptional read-through to an extent, but not completely. The basic PhlF gates tested in the literature [100], use the exact same terminator (ECK120033737) as we used to terminate transcription of the PhlF, but they do not see signs of transcriptional read-through. Transcription of PhlF in the basic PhlF gate was regulated by P_{tac} but in our tuneable gate it was regulated by P_{tet}, which has a maximum promoter activity 37% higher than P_{tac}[100]. Therefore, in our system, when the input promoter activity is at its highest, 37% more RNAPs are recruited to transcribe PhlF, so more RNAPs would read through the terminator than in the basic gate. This could account for the difference in output promoter activity at high input promoter activity.

These results highlight the importance of using more insightful and quantitative methods of capturing how well parts work, such as RNA sequencing [50]. In the case of our tuneable NOT gate, transcriptional read-through appears to reduce the devices performance by limiting the fold-change in output promoter activities when input promoter activities are low and high.

Even if the performance of genetic devices is improved by building them in a way that prevents transcriptional read-through, the environment that cells are harbouring a device are grown in can impact performance. If circuits are not being tested in environments relevant to the conditions they're built to be used in, we cannot predict how they will behave in industry, making it difficult to design and build circuits for this purpose.

To emulate environmental conditions that a device would be exposed to in industry, we grew cells in 0.04 , 0.4 and 4 g L^{-1} of glucose and measured their fluorescence at 4 time points over 180 minutes, as cells were grown in varying input and tuner promoter activities. The fold-change in the device's protein production rate when it was grown with high and low input promoter activities, was 10% and 25% lower at low and high tuner promoter activities, respectively, when cells grown were grown in 4 g L^{-1} glucose compared to cells grown in 0.04 g L^{-1} glucose. This is likely because growing cells in low glucose concentrations stresses them, causing them to express a number of genes that alter their physiology in a way that further reduces the rate they produce protein [75, 75, 121].

We would expect cells grown in the highest concentration of glucose (4 g L^{-1}), to produce protein at a greater rate than cells grown in 0.4 g L^{-1}, but they don't. If cells are grown in more glucose than they need to survive, they grow too fast, using up oxygen available to them. Cells starved of oxygen use different native metabolic pathways which changes the availability of cellular resources and, therefore, the rate of protein production [42]. This could explain why, despite the growth rate of cells increasing at 4 g L^{-1} compared to 0.4 g L^{-1}, the rate of production does not increase. To validate this hypothesis, we could grow cells in media containing varying glucose concentrations and measure the dissolved oxygen concentration in the media.

Finally, we show that for a fixed input promoter activity of 6.06 RPU, the TES can be tuned so that the rate of protein production is always the same, regardless of changes in glucose concentration. This result is significant as it shows that the TES could be tuned to maintain constant performance as changes in the environment change rates of gene expression. Previous efforts have been made to build circuits that dynamically adapt to their environment, by using sensors that detect the concentration of some external chemical [96] or by activating circuits in response to an internal cellular process, such as metabolic burden [27]. Our devices could be combined with such approaches, so that the performance of a device or circuit could be dynamically tuned in response to changes in external chemical concentration or metabolic burden. Furthermore, our tuneable devices offer synthetic biologists another tool they can incorporate into circuits, which could reduce the time taken to engineer a working product.

Previously when circuits broke due to imbalanced gene expression, or mismatched device response functions, they would have to be redesigned and rebuilt. Our tuneable devices provide an alternative approach. They can be incorporated into larger circuits and if the circuit breaks, instead of going through the time consuming process of re-assembling a circuit, it can be fixed by tuning the devices.

CHAPTER 7

Conclusions

The motivation behind this theses was to to develop tools and methodologies that would help reduce the number of DBTL cycles needed to engineer working genetic systems. We aimed to build genetic devices whose behaviours can be dynamically modified after their assembly, cutting the cost and time needed to engineer genetic circuits. If a circuit built from our devices does not work as expected, we can simply tune the devices to fix the system.

To tackle this problem, we developed a tuneable expression system (TES), where the rate of transcription and translation of a gene of interest can be dynamically controlled. We optimized the device's performance by improving its fold change and usability in populations of cells and show that it can be used to build a tuneable genetic NOT gate. Following on from this we showed that changing the glucose concentration that cells are grown in alters the rate that they produce protein. Finally, we demonstrated that the TES can tune protein production rates so that they remain constant in different concentrations of glucose.

In Chapter 4, we designed, built and tested a prototype TES. We modeled the TES's behaviour and showed how the relationship between the input and output can be modified. To realize the device *in-vivo*, genetic sensors P_{tet} and P_{tac} were used to regulate its input and tuner. We used the TES to vary the rate of protein production over a range of inputs and found that we could alter the shape of its response function. To investigate the dynamics of protein production by the TES, we characterised protein production rate over time when the device's inputs were varied.

Whilst our initial design demonstrated that the concept of a TES is feasible, the device didn't perform well. For a signal to propagate between connected devices, the range of outputs from the first device must be large enough to span the inputs to the second [13, 130]. We expected the range of outputs from the TES to be similar to those of the input sensor, but they're significantly smaller. The limited range of outputs from the device make it incompatible for use in large circuits, where devices' outputs need to have large dynamic ranges for a signal to propagate robustly [100]. Furthermore, individual cells in populations grown at low and high input promoter

activities showed similar output fluorescences, so for a large percentage of cells in any population, it is impossible to determine whether the device's output is low or high. For large genetic circuits to function as intended, they need to be built from devices that produce two distinct signals when their inputs are in low and high states [6, 100].

To address this limitation, in Chapter 5, we aimed to optimize the TES design. We used deterministic and stochastic models to show that increasing the rate that sRNA is transcribed increases the device's fold change and decreases the intersection between fluorescence distributions for the TES. We built a second "sRNA booster" plasmid that lead to overall higher transcription rates for the sRNA than the original device when co-transformed with the TES. We characterised the effect of the sRNA booster plasmid and confirmed what our model had predicted. In our initial design we used an insulator part that might have interfered with how the THS RNA folds and functions, hindering the TESs performance. A thermodynamic model showed that the insulator physically binds to the THS and decreases the rate that the THS folds, preventing repression of translation and sRNA binding to the THS to activate translation. To test this *in-vivo*, we constructed a non-insulated TES where the insulator part was removed. As predicted, removing the insulator caused an increase in the fold-change of device outputs at all tuner promoter activities. We combined the non-insulated TES with the sRNA booster and found that increasing the rate that sRNA is transcribed for a non-insulated TES further increased the fold-change and decreased the intersection between populations of cells. The RiboJ insulator that we used in the TES is routinely used to make genetic device behaviour more predictable [34, 83, 100]. We showed that the insulator can interfere with how RNA based parts function. If synthetic biologists are using ribozyme based insulators directly upstream of an RNA part, our results recommend the use of thermodynamic modelling to check for possible interference.

Despite all these improvements, the fold-change of the TES still decreased substantially as we tuned it. A potential explanation is that when we increase the rate of transcription of sRNA to tune our device, RNAP reads through the terminator and transcribes the downstream gene, producing mRNA even when the input promoter is not active [18, 32]. We chose strong terminators in all our TES designs [32]. However, our results suggest that the L3S3P11 terminator we use to terminate transcription of the sRNA was ineffective. This could be down to the way the terminator was characterised. For example, the fluorescent protein based method used to characterise terminators is not robust to contextual effects, so it is possible that the terminators are not as effective when used in our circuit.

To connect genetic devices, their inputs and outputs need to use the same signals [19, 100]. In Chapter 6 we designed and built a tuneable NOT gate, whose input and output is RNAP flux at a promoter, by replacing the YFP in the TES design with the repressor protein PhlF and adding a new output promoter, P_{phlF}. We built NOT gates for the four variants of the TES: the original design, the original design with a booster, the non-insulated design and the non-insulated design with the booster. We showed that we can tune the transition point of all the gates. NOT gates have

been shown to be valuable devices for building large circuits that carry out complicated functions, however, large circuits rarely function as intended [100, 124]. One common reason circuits break is that the input signal cannot propagate through the circuit due to the transition points of connected devices being mismatched [13, 100, 130]. Until now, tuning a device's transition point, required rebuilding the circuit. In contrast, the tuneable NOT gate could be integrated into such circuits and used to tune and fix transition points of mismatched devices; a quicker and cheaper way to tune and fix a circuit. Tuneable NOT gates that can have their transition points tuned over a larger range embody the functionality of a larger range of different "static" NOT gates. They are better suited to be used within larger circuits, as circuits built with such a device encompass a larger design space within a fragment of DNA of similar length. To this end, the gate built using the original TES design performed best as it could be tuned over the largest range of transition points and so embodied the functionality of the largest range of NOT gates.

Genetic circuits can be used to engineer cells to produce chemicals in industrial bioreactors. However, bioreactors expose cells to a large range of different environmental conditions (i.e. fluctuations in glucose concentration [42]), which may affect how a genetic circuit performs [52, 95]. To overcome this problem, we show that the TES could be used to adapt protein production rates so that they are constant in a range of glucose concentrations. By characterising the TES in different concentrations of glucose, we demonstrated that the TES performs worse at low glucose concentrations, when cells grow slower. In these conditions proteins were produced at a lower rate which impacted the device's performance. These results are interesting for anyone engineering genetic systems for use in industrial bioreactors, where cells experience different concentrations of glucose depending on where in the reactor they are at different points in time. For genetic circuits that need to produce precise concentrations of proteins to function, this would cause them to break [2, 138]. We show that for genes regulated by the TES, the rate of protein production can be dynamically tuned so that cells produce protein at a constant rate in different concentrations of glucose.

TESs could be valuable for engineering adaptive systems that tune gene expression in response to changes in their environment. However, uses of the TESs we present here are limited as they rely on external inputs (small molecules aTc and IPTG) to regulate gene expression. For example, if we wanted to regulate expression of a gene in an industrial context where cells are grown in large volumes of media, it would be too expensive and impractical to rely on such inputs. Therefore, for the TES to be of practical use new autonomous TESs need to be designed to use inputs and tuner inputs that regulate gene expression in response to signals that are either found in the environment they are to be used in, such as substrate concentration or pH [77], or signals that occur within cells, such as responses to metabolic burden [27].

The findings might also be of use to chemical engineers looking to find optimal concentrations of glucose to grow cells in for producing proteins. The method we use, growing cells in different conditions in parallel, can be applied to test circuits in ranges of other conditions relevant to

industry. This work provides a first step towards a pipeline that provides a lab-scale environment for testing genetic circuits and scaling their use for industrial contexts.

More broadly, this work contributes to the field of synthetic biology in a number of ways. First, we have generated two sets of devices. (i) TESs that allow the rate of protein production to be tuned and (ii) tunebable NOT gates that can have their transition points between on and off states tuned and that are compatible for integration into large circuits that carry out more complex functions [100]. Secondly, we discovered an important design rule for those using Ribozyme-based insulators: when using insulators adjacent to RNA devices, ensure that their secondary structures don't interfere with each others function. Finally, we showed that the performance of our devices changes in different concentrations of glucose. We then demonstrated that we can tune the devices to produce protein at a constant rate in varying glucose concentrations, which will be of interest to parties looking to build genetic circuits that are robust to perturbations in environmental conditions.

7.1 Future directions

7.1.1 Use in biodesign automation

The contents of this thesis are useful to other researchers, however, there are several limitations in what we can infer from the results we generate. We state that the way we designed a tuneable NOT gate in Chapter 6 would allow us to use the device in biodesign automation (BDA) software such as Cello [101]. Despite this being theoretically possible, we never tested out how well the device would work in the software or whether it can be used to tune and fix larger circuits. A future direction for this work would be to integrate the NOT gate design into Cello and use the software to design a large genetic circuit. The circuits performance can then be improved by tuning NOT gates.

7.1.2 Automated characterisation of devices

To build on our initial characterisations we could automate the characterisation of our devices to yield large amounts of data that's more reproducible than if experiments were done by hand [20]. Furthermore, automated sampling of growing cell cultures could generate high resolution time course data, with which we could more accurately fit model parameters, yielding models that better describe our experimental system and allow us to more accurately predict device behaviour *in-silico*. Furthermore, automation would enable us to characterise our tuneable devices in a large number of different environmental conditions in parallel, offering insight into how they are best used.

7.1.3 Testing devices in different environments

In Chapter 6 we showed that testing the TES in different concentrations of glucose affects its performance. Environmental conditions change how genetic parts and devices that make up a circuit perform, which can cause the circuit to break. For genetic circuits to be used in real world applications, where they are exposed to a range of environmental conditions, their performance must be robust to the conditions they experience. Our tuneable devices allow us to address this problem. By tuning the rate of protein production for individual genes in a circuit, we can dynamically tune protein production rates so they're constant in any environment, making circuits adaptable to changing conditions. The next step would be to characterise our devices in a range of conditions that engineered cells would be exposed to in industry, such as: dissolved oxygen concentrations, available nitrogen (i.e. amino acid) concentrations, pHs and temperatures. From here, we could swap the tuner input for a genetic sensor that detects changes in a condition of interest, allowing the device to automatically tune its response function in response to environmental perturbations.

7.1.4 Dynamically fixing broken circuits

We showed that we can build a tuneable NOT gate and that we can tune the gate's transition point between on and off states. A natural next step would be to build dynamic circuits that tune themselves in response to changes in host cell physiology or their environment, making their function robust to changes in contexts that would otherwise break them. For example, if a cell produces a toxic protein, we could regulate the transcription of the protein using a tuneable NOT gate. The THS part of the device could be coupled to an internal cellular process that tunes the gate when the cell is stressed, down regulating the production of the toxic protein. An alarmone such as ppGpp could be used as an indicator that cells are stressed. Instead of using a THS, we could use a ppGpp responsive riboswitch to regulate the production of the repressor protein [120].

7.2 Outlook

The work presented here is a first step towards a new, more effective approach to engineering living systems. Using our tuneable devices or taking inspiration from them, synthetic biologists can start to move away from the design-build-test-learn cycle, in favour of a more intelligent approach, where they can side-step days or even months of painstaking DNA assembly using our novel design-build-test-tune framework. If widely adopted, these innovations could accelerate the rate that functioning genetic circuits are engineered and in doing so advance the rate that world changing biotechnologies are deployed into society.

APPENDIX A

Appendix A

A.1 Plasmid Sequences

A.1.1 pVB001

5'--GCTTCCTCGCTCACTGACTCGCTGCACGAGGCAGACCTCAGCGCTAGCGGAGTGTATA
CTGGCTTACTATGTTGGCACTGATGAGGGTGTCAGTGAAGTGCTTCATGTGGCAGGAGAA
AAAAGGCTGCACCGGTGCGTCAGCAGAATATGTGATACAGGATATATTCCGCTTCCTCGCT
CACTGACTCGCTACGCTCGGTCGTTCGACTGCGGCGAGCGGAAATGGCTTACGAACGGGG
CGGAGATTTCCCGGAAGATGCCAGGAAGATACTTAACAGGGAAGTGAGAGGGCCGCGGC
AAAGCCGTTTTTCCATAGGCTCCGCCCCCCTGACAAGCATCACGAAATCTGACGCTCAAAT
CAGTGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGCGGCTCC
CTCGTGCGCTCTCCTGTTCCTGCCTTTCGGTTTACCGGTGTCATTCCGCTGTTATGGCCGC
GTTTGTCTCATTCCACGCCTGACACTCAGTTCCGGGTAGGCAGTTCGCTCCAAGCTGGACT
GTATGCACGAACCCCCCGTTCAGTCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGA
GTCCAACCCGGAAAGACATGCAAAAGCACCACTGGCAGCAGCCACTGGTAATTGATTTAG
AGGAGTTAGTCTTGAAGTCATGCGCCGGTTAAGGCTAAACTGAAAGGACAAGTTTTGGTG
ACTGCGCTCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTAGCTCAGAGAACCTTCGA
AAAACCGCCCTGCAAGGCGGTTTTTTCGTTTTCAGAGCAAGAGATTACGCGCAGACCAAA
ACGATCTCAAGAAGATCATCTTATTAAGGGGTCTGACGCTCAGTGGAACGAAAAATCAATC
TAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCTTAGAAAAACTCATCGAGCATCAA
ATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCT
GTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGT
CTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAG
GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGCTT
ATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTC
GCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCG

CTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGC
GCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCC
CGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGG
TCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATT
GGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAT
CGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAAT
CAGCATCCATGTTGGAATTTAATCGCGGCCTGGAGCAAGACGTTTCCCGTTGAATATGGCT
CATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATAT
TTTTATCTTGTGCAATGTACATCAGAGATTTTGAGACACAACCAATTATTGAAGGCCTCCC
TAACGGGGGGCCTTTTTTTGAACGATCGTTGGCTGTGTTGACAATTAATCATCGGCTCGTA
TAATGTGTGGAATTGTGAGCGCTCACAATTGGGTCATGACTGGGACACGCCAGTCATGAG
AATACAGACATAAAGCAAGCCAGAGATTAACGAAGCCAATTATTGAACACCCTTCGGGGT
GTTTTTTTGTTTCTGGTCTACCTACTCCACCGTTGGCTTTTTTCCCTATCAGTGATAGAGAT
TGACATCCCTATCAGTGATAGAGATAATGAGCACAGCTGTCACCGGATGTGCTTTCCGGTC
TGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAGGGCGTTAATCTC
TGGCTTGCTTTATGTCTGTAAACAGAGGAGATACAGAATGAAAGCAAGCAACCTGGCGGC
AGCGCAAAAGATGCGTAAAGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCA
TCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGC
GAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACAGGCAAGCTG
CCCGTGCCCTGGCCCACCCTCGTGACCACCTTCGGCTACGGCCTGCAATGCTTCGCCCGC
TACCCCGACCACATGAAGCTGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTC
CAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAG
TTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGAC
GGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATG
GCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGAC
GGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCAATCGGCGACGGCCCCGT
GCTGCTGCCCGACAACCACTACCTTAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGA
GAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCAT
GGACGAGCTGTACAAGTAACTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGG
GCCTTTTTTCGTTTTGGTCCAATGGTCACCATATATCAAGTTTACGGCTAGCTCAGTCCTA
GGTACTATGCTAGCTACTAGAGAAAGAGGAGAAATACTAGATGGCTGAAGCGCAAAATGA
TCCCCTGCTGCCGGGATACTCGTTTAATGCCCATCTGGTGGCGGGTTTAACGCCGATTGA
GGCCAACGGTTATCTCGATTTTTTTATCGACCGACCGCTGGGAATGAAAGGTTATATTCTC
AATCTCACCATTCGCGGTCAGGGGGTGGTGAAAAATCAGGGACGAGAATTTGTTTGCCGA
CCGGGTGATATTTTGCTGTTCCCGCCAGGAGAGATTCATCACTACGGTCGTCATCCGGAG
GCTCGCGAATGGTATCACCAGTGGGTTTACTTTCGTCCGCGCGCCTACTGGCATGAATGG
CTTAACTGGCCGTCAATATTTGCCAATACGGGGTTCTTTCGCCCGGATGAAGCGCACCAG
CCGCATTTCAGCGACCTGTTTGGGCAAATCATTAACGCCGGGCAAGGGGAAGGGCGCTAT

TCGGAGCTGCTGGCGATAAATCTGCTTGAGCAATTGTTACTGCGGCGCATGGAAGCGATT
AACGAGTCGCTCCATCCACCGATGGATAATCGGGTACGCGAGGCTTGTCAGTACATCAGC
GATCACCTGGCAGACAGCAATTTTGATATCGCCAGCGTCGCACAGCATGTTTGCTTGTCG
CCGTCGCGTCTGTCACATCTTTTCCGCCAGCAGTTAGGGATTAGCGTCTTAAGCTGGCGC
GAGGACCAACGTATCAGCCAGGCGAAGCTGCTTTTGAGCACCACCCGGATGCCTATCGCC
ACCGTCGGTCGCAATGTTGGTTTTGACGATCAACTCTATTTCTCGCGGGTATTTAAAAAAT
GCACCGGGGCCAGCCCGAGCGAGTTCCGTGCCGGTTAATAACCAATTATTGAAGGCCGCT
AACGCAGCCTTTTTTTGTTTCTGGTCTCCCAATGGCGGCGCGCCATCGAATGGCGCAAAAC
CTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATATGAA
ACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGC
GTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGAT
GGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTT
GCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGC
GATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAG
CGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCT
GATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAAT
GTTCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCC
ATGAGGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCG
CGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATA
AATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCA
TGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCT
GGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGC
GCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGATAGCTCATGTTATA
TCCCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACC
GCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCAGTCTCAC
TGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGG
CCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGACAAT
CCAGGAGGAAAAAAATGTCCAGATTAGATAAAAGTAAAGTGATTAACAGCGCATTAGAGC
TGCTTAATGAGGTCGGAATCGAAGGTTTAACAACCCGTAAACTCGCCCAGAAGCTAGGTG
TAGAGCAGCCTACATTGTATTGGCATGTAAAAAATAAGCGGGCTTTGCTCGACGCCTTAGC
CATTGAGATGTTAGATAGGCACCATACTCACTTTTGCCCTTTAGAAGGGGAAAGCTGGCAA
GATTTTTTACGTAATAACGCTAAAAGTTTTAGATGTGCTTTACTAAGTCATCGCGATGGAG
CAAAAGTACATTTAGGTACACGGCCTACAGAAAAACAGTATGAAACTCTCGAAAATCAATT
AGCCTTTTTATGCCAACAAGGTTTTTCACTAGAGAATGCATTATATGCACTCAGCGCTGTG
GGGCATTTTACTTTAGGTTGCGTATTGGAAGATCAAGAGCATCAAGTCGCTAAAGAAGAA
AGGGAAACACCTACTACTGATAGTATGCCGCCATTATTACGACAAGCTATCGAATTATTTG
ATCACCAAGGTGCAGAGCCAGCCTTCTTATTCGGCCTTGAATTGATCATATGCGGATTAGA
AAAACAACTTAAATGTGAAAGTGGGTCCTAATAATTGGTAACGAATCAGACAATTGACGGC

TCGAGGGAGTAGCATAGGGTTTGCAGAATCCCTGCTTCGTCCATTTGACAGGCACATTAT
GCATCGATGATAAGCTGTCAAACATGAGCAGATCCTCTACGCCGGACGCATCGTGGCCGG
CATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGA
AGATCGGGCTCGCCACTTCGGGCTCATGAGCAAATATTTTATCTGAGGT--3’

A.1.2 pVB002

5’--GCTTCCTCGCTCACTGACTCGCTGCACGAGGCAGACCTCAGCGCTAGCGGAGTGTATA
CTGGCTTACTATGTTGGCACTGATGAGGGTGTCAGTGAAGTGCTTCATGTGGCAGGAGAA
AAAAGGCTGCACCGGTGCGTCAGCAGAATATGTGATACAGGATATATTCCGCTTCCTCGCT
CACTGACTCGCTACGCTCGGTCGTTCGACTGCGGCGAGCGGAAATGGCTTACGAACGGGG
CGGAGATTTCCTGGAAGATGCCAGGAAGATACTTAACAGGGAAGTGAGAGGGCCGCGGC
AAAGCCGTTTTTCCATAGGCTCCGCCCCCCTGACAAGCATCACGAAATCTGACGCTCAAAT
CAGTGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGCGGCTCC
CTCGTGCGCTCTCCTGTTCCTGCCTTTCGGTTTACCGGTGTCATTCCGCTGTTATGGCCGC
GTTTGTCTCATTCCACGCCTGACACTCAGTTCCGGGTAGGCAGTTCGCTCCAAGCTGGACT
GTATGCACGAACCCCCCGTTCAGTCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGA
GTCCAACCCGGAAAGACATGCAAAAGCACCACTGGCAGCAGCCACTGGTAATTGATTTAG
AGGAGTTAGTCTTGAAGTCATGCGCCGGTTAAGGCTAAACTGAAAGGACAAGTTTTGGTG
ACTGCGCTCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTAGCTCAGAGAACCTTCGA
AAAACCGCCCTGCAAGGCGGTTTTTTCGTTTTCAGAGCAAGAGATTACGCGCAGACCAAA
ACGATCTCAAGAAGATCATCTTATTAAGGGGTCTGACGCTCAGTGGAACGAAAAATCAATC
TAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCTTAGAAAAACTCATCGAGCATCAA
ATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCT
GTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGT
CTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAG
GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATAGCAAAAGCTT
ATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTC
GCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCG
CTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGC
GCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCC
CGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGG
TCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATT
GGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAT
CGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAAT
CAGCATCCATGTTGGAATTTAATCGCGGCCTGGAGCAAGACGTTTCCCGTTGAATATGGCT
CATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATAT
TTTTATCTTGTGCAATGTACATCAGAGATTTTGAGACACAACCAATTATTGAAGGCCTCCC
TAACGGGGGGCCTTTTTTTGAACGATCGTTGGCTGTGTTGACAATTAATCATCGGCTCGTA

TAATGTGTGGAATTGTGAGCGCTCACAATTGGGTCATGACTGGGACACGCCAGTCATGAG
AATACAGACATAAAGCAAGCCAGAGATTAACGAAGCCAATTATTGAACACCCTTCGGGGT
GTTTTTTTGTTTCTGGTCTACCTACTCCACCGTTGGCTTTTTTCCCTATCAGTGATAGAGAT
TGACATCCCTATCAGTGATAGAGATAATGAGCACAGCTGTCACCGGATGTGCTTTCCGGTC
TGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAAGGGCGTTAATCTC
TGGCTTGCTTTATGTCTGTAAACAGAGGAGATACAGAATGAAAGCAAGCAACCTGGCGGC
AGCGCAAAAGATGCGTAAAGCACGTACCCCGAGCCGTAGCAGCATTGGTAGCCTGCGTAG
TCCGCATACCCATAAAGCAATTCTGACCAGCACCATTGAAATCCTGAAAGAATGTGGTTAT
AGCGGTCTGAGCATTGAAAGCGTTGCACGTCGTGCCGGTGCAAGCAAACCGACCATTTAT
CGTTGGTGGACCAATAAAGCAGCACTGATTGCCGAAGTGTATGAAAATGAAAGCGAACAG
GTGCGTAAATTTCCGGATCTGGGTAGCTTTAAAGCCGATCTGGATTTTCTGCTGCGTAATC
TGTGGAAAGTTTGGCGTGAAACCATTTGTGGTGAAGCATTTCGTTGTGTTATTGCAGAAGC
ACAGCTGGACCCTGCAACCCTGACCCAGCTGAAAGATCAGTTTATGGAACGTCGTCGTGA
GATGCCGAAAAAACTGGTTGAAAATGCCATTAGCAATGGTGAACTGCCGAAAGATACCAA
TCGTGAACTGCTGCTGGATATGATTTTTGGTTTTTGTTGGTATCGCCTGCTGACCGAACAG
CTGACCGTTGAACAGGATATTGAAGAATTTACCTTCCTGCTGATTAATGGTGTTTGTCCGG
GTACACAGCGTTAAGGAAACACAGAAAAAAGCCCGCACCTGACAGTGCGGGCTTTTTTTT
TCGACCAAAGGTGTCAACGTTCGACGTACGGTGGAATCTGATTCGTTACCAATTGACATGA
TACGAAACGTACCGTATCGTTAAGGTCAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGA
GTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAATACTAGAGAAAGAGGGGA
AATACTAGATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCG
AGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGAT
GCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACAGGCAAGCTGCCCGTGCCC
TGGCCCACCCTCGTGACCACCTTCGGCTACGGCCTGCAATGCTTCGCCCGCTACCCCGAC
CACATGAAGCTGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGC
ACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGC
GACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATC
CTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAG
CAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTG
CAGCTCGCCGACCACTACCAGCAGAACACCCCAATCGGCGACGGCCCCGTGCTGCTGCCC
GACAACCACTACCTTAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGAT
CACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTG
TACAAGTAACTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTC
GTTTTGGTCCAATGGCGGCGCGCCATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGA
TAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATATGAAACCAGTAACGTTATACGAT
GTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGC
CACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATT
CCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACC

TCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGAT
CAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAA
GCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTG
GATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTG
ATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAGGACGGTACGCGACT
GGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATT
AAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAA
ATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACC
ATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATG
GCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTCG
GTAGTGGGATACGACGATACCGAAGATAGCTCATGTTATATCCCGCCGTTAACCACCATCA
AACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGG
GCCAGGCGGTGAAGGGCAATCAGCTGTTGCCAGTCTCACTGGTGAAAAGAAAAACCACCC
TGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGG
CACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGATAATCCAGGAGGAAAAAAATGTCCA
GATTAGATAAAAGTAAAGTGATTAACAGCGCATTAGAGCTGCTTAATGAGGTCGGAATCG
AAGGTTTAACAACCCGTAAACTCGCCCAGAAGCTAGGTGTAGAGCAGCCTACATTGTATT
GGCATGTAAAAAATAAGCGGGCTTTGCTCGACGCCTTAGCCATTGAGATGTTAGATAGGC
ACCATACTCACTTTTGCCCTTTAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATAACGC
TAAAAGTTTTAGATGTGCTTTACTAAGTCATCGCGATGGAGCAAAAGTACATTTAGGTACA
CGGCCTACAGAAAAACAGTATGAAACTCTCGAAAATCAATTAGCCTTTTTATGCCAACAAG
GTTTTTCACTAGAGAATGCATTATATGCACTCAGCGCTGTGGGGCATTTTACTTTAGGTTG
CGTATTGGAAGATCAAGAGCATCAAGTCGCTAAAGAAGAAAGGGAAACACCTACTACTGA
TAGTATGCCGCCATTATTACGACAAGCTATCGAATTATTTGATCACCAAGGTGCAGAGCCA
GCCTTCTTATTCGGCCTTGAATTGATCATATGCGGATTAGAAAAACAACTTAAATGTGAAA
GTGGGTCCTAATAATTGGTAACGAATCAGACAATTGACGGCTCGAGGGAGTAGCATAGGG
TTTGCAGAATCCCTGCTTCGTCCATTTGACAGGCACATTATGCATCGATGATAAGCTGTCA
AACATGAGCAGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTG
CGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCG
GGCTCATGAGCAAATATTTTATCTGAGGT--3’

A.1.3 pVB003

5’--GCTTCCTCGCTCACTGACTCGCTGCACGAGGCAGACCTCAGCGCTAGCGGAGTGTATA
CTGGCTTACTATGTTGGCACTGATGAGGGTGTCAGTGAAGTGCTTCATGTGGCAGGAGAA
AAAAGGCTGCACCGGTGCGTCAGCAGAATATGTGATACAGGATATATTCCGCTTCCTCGCT
CACTGACTCGCTACGCTCGGTCGTTCGACTGCGGCGAGCGGAAATGGCTTACGAACGGGG
CGGAGATTTCCCGGAAGATGCCAGGAAGATACTTAACAGGGAAGTGAGAGGGCCGCGGC
AAAGCCGTTTTTCCATAGGCTCCGCCCCCCTGACAAGCATCACGAAATCTGACGCTCAAAT

CAGTGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGCGGCTCC
CTCGTGCGCTCTCCTGTTCCTGCCTTTCGGTTTACCGGTGTCATTCCGCTGTTATGGCCGC
GTTTGTCTCATTCCACGCCTGACACTCAGTTCCGGGTAGGCAGTTCGCTCCAAGCTGGACT
GTATGCACGAACCCCCCGTTCAGTCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGA
GTCCAACCCGGAAAGACATGCAAAAGCACCACTGGCAGCAGCCACTGGTAATTGATTTAG
AGGAGTTAGTCTTGAAGTCATGCGCCGGTTAAGGCTAAACTGAAAGGACAAGTTTTGGTG
ACTGCGCTCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTAGCTCAGAGAACCTTCGA
AAAACCGCCCTGCAAGGCGGTTTTTTCGTTTTCAGAGCAAGAGATTACGCGCAGACCAAA
ACGATCTCAAGAAGATCATCTTATTAAGGGGTCTGACGCTCAGTGGAACGAAAAATCAATC
TAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCTTAGAAAAACTCATCGAGCATCAA
ATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCT
GTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGT
CTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAG
GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGCTT
ATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTC
GCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCG
CTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGC
GCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCC
CGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGG
TCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATT
GGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAT
CGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAAT
CAGCATCCATGTTGGAATTTAATCGCGGCCTGGAGCAAGACGTTTCCCGTTGAATATGGCT
CATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATAT
TTTTATCTTGTGCAATGTACATCAGAGATTTTGAGACACAACCAATTATTGAAGGCCTCCC
TAACGGGGGGCCTTTTTTTGAACGATCGTTGGCTGTGTTGACAATTAATCATCGGCTCGTA
TAATGTGTGGAATTGTGAGCGCTCACAATTGGGTCATGACTGGGACACGCCAGTCATGAG
AATACAGACATAAAGCAAGCCAGAGATTAACGAAGCCAATTATTGAACACCCTTCGGGGT
GTTTTTTTGTTTCTGGTCTACCTACTCCACCGTTGGCTTTTTTCCCTATCAGTGATAGAGAT
TGACATCCCTATCAGTGATAGAGATAATGAGCACGGGCGTTAATCTCTGGCTTGCTTTATG
TCTGTAAACAGAGGAGATACAGAATGAAAGCAAGCAACCTGGCGGCAGCGCAAAAGATGC
GTAAAGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTG
GACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCAC
CTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACAGGCAAGCTGCCCGTGCCCTGGCC
CACCCTCGTGACCACCTTCGGCTACGGCCTGCAATGCTTCGCCCGCTACCCCGACCACAT
GAAGCTGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCAT
CTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACAC
CCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGG

GCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA
GAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT
CGCCGACCACTACCAGCAGAACACCCCAATCGGCGACGGCCCCGTGCTGCTGCCCGACAA
CCACTACCTTAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACAT
GGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAA
GTAACTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTTCGTTTT
GGTCCAATGGTCACCATATATCAAGTTTACGGCTAGCTCAGTCCTAGGTACTATGCTAGCT
ACTAGAGAAAGAGGAGAAATACTAGATGGCTGAAGCGCAAAATGATCCCCTGCTGCCGGG
ATACTCGTTTAATGCCCATCTGGTGGCGGGTTTAACGCCGATTGAGGCCAACGGTTATCTC
GATTTTTTTATCGACCGACCGCTGGGAATGAAAGGTTATATTCTCAATCTCACCATTCGCG
GTCAGGGGGTGGTGAAAAATCAGGGACGAGAATTTGTTTGCCGACCGGGTGATATTTTGC
TGTTCCCGCCAGGAGAGATTCATCACTACGGTCGTCATCCGGAGGCTCGCGAATGGTATC
ACCAGTGGGTTTACTTTCGTCCGCGCGCCTACTGGCATGAATGGCTTAACTGGCCGTCAAT
ATTTGCCAATACGGGGTTCTTTCGCCCGGATGAAGCGCACCAGCCGCATTTCAGCGACCT
GTTTGGGCAAATCATTAACGCCGGGCAAGGGGAAGGGCGCTATTCGGAGCTGCTGGCGAT
AAATCTGCTTGAGCAATTGTTACTGCGGCGCATGGAAGCGATTAACGAGTCGCTCCATCCA
CCGATGGATAATCGGGTACGCGAGGCTTGTCAGTACATCAGCGATCACCTGGCAGACAGC
AATTTTGATATCGCCAGCGTCGCACAGCATGTTTGCTTGTCGCCGTCGCGTCTGTCACATC
TTTTCCGCCAGCAGTTAGGGATTAGCGTCTTAAGCTGGCGCGAGGACCAACGTATCAGCC
AGGCGAAGCTGCTTTTGAGCACCACCCGGATGCCTATCGCCACCGTCGGTCGCAATGTTG
GTTTTGACGATCAACTCTATTTCTCGCGGGTATTTAAAAAATGCACCGGGGCCAGCCCGA
GCGAGTTCCGTGCCGGTTAATAACCAATTATTGAAGGCCGCTAACGCAGCCTTTTTTTGTT
TCTGGTCTCCCAATGGCGGCGCGCCATCGAATGGCGCAAAACCTTTCGCGGTATGGCATG
ATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATATGAAACCAGTAACGTTATACGA
TGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAG
CCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACAT
TCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCAC
CTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGA
TCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAA
AGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCT
GGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTT
GATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAGGACGGTACGCGAC
TGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCAT
TAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCA
AATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAAC
CATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGAT
GGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTC
GGTAGTGGGATACGACGATACCGAAGATAGCTCATGTTATATCCCGCCGTTAACCACCAT

CAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCA
GGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCAGTCTCACTGGTGAAAAGAAAAACCAC
CCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCT
GGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGACAATCCAGGAGGAAAAAAATGTC
CAGATTAGATAAAAGTAAAGTGATTAACAGCGCATTAGAGCTGCTTAATGAGGTCGGAAT
CGAAGGTTTAACAACCCGTAAACTCGCCCAGAAGCTAGGTGTAGAGCAGCCTACATTGTA
TTGGCATGTAAAAAATAAGCGGGCTTTGCTCGACGCCTTAGCCATTGAGATGTTAGATAG
GCACCATACTCACTTTTGCCCTTTAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATAAC
GCTAAAAGTTTTAGATGTGCTTTACTAAGTCATCGCGATGGAGCAAAAGTACATTTAGGTA
CACGGCCTACAGAAAAACAGTATGAAACTCTCGAAAATCAATTAGCCTTTTTATGCCAACA
AGGTTTTTCACTAGAGAATGCATTATATGCACTCAGCGCTGTGGGGCATTTTACTTTAGGT
TGCGTATTGGAAGATCAAGAGCATCAAGTCGCTAAAGAAGAAAGGGAAACACCTACTACT
GATAGTATGCCGCCATTATTACGACAAGCTATCGAATTATTTGATCACCAAGGTGCAGAGC
CAGCCTTCTTATTCGGCCTTGAATTGATCATATGCGGATTAGAAAAACAACTTAAATGTGA
AAGTGGGTCCTAATAATTGGTAACGAATCAGACAATTGACGGCTCGAGGGAGTAGCATAG
GGTTTGCAGAATCCCTGCTTCGTCCATTTGACAGGCACATTATGCATCGATGATAAGCTGT
CAAACATGAGCAGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGG
TGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTT
CGGGCTCATGAGCAAATATTTTATCTGAGGT--3’

A.1.4 pVB004

5’--GCTTCCTCGCTCACTGACTCGCTGCACGAGGCAGACCTCAGCGCTAGCGGAGTGTATA
CTGGCTTACTATGTTGGCACTGATGAGGGTGTCAGTGAAGTGCTTCATGTGGCAGGAGAA
AAAAGGCTGCACCGGTGCGTCAGCAGAATATGTGATACAGGATATATTCCGCTTCCTCGCT
CACTGACTCGCTACGCTCGGTCGTTCGACTGCGGCGAGCGGAAATGGCTTACGAACGGGG
CGGAGATTTCCTGGAAGATGCCAGGAAGATACTTAACAGGGAAGTGAGAGGGCCGCGGC
AAAGCCGTTTTTCCATAGGCTCCGCCCCCCTGACAAGCATCACGAAATCTGACGCTCAAAT
CAGTGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGCGGCTCC
CTCGTGCGCTCTCCTGTTCCTGCCTTTCGGTTTACCGGTGTCATTCCGCTGTTATGGCCGC
GTTTGTCTCATTCCACGCCTGACACTCAGTTCCGGGTAGGCAGTTCGCTCCAAGCTGGACT
GTATGCACGAACCCCCCGTTCAGTCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGA
GTCCAACCCGGAAAGACATGCAAAAGCACCACTGGCAGCAGCCACTGGTAATTGATTTAG
AGGAGTTAGTCTTGAAGTCATGCGCCGGTTAAGGCTAAACTGAAAGGACAAGTTTTGGTG
ACTGCGCTCCTCCAAGCCAGTTACCTCGGTTCAAAGAGTTGGTAGCTCAGAGAACCTTCGA
AAAACCGCCCTGCAAGGCGGTTTTTTCGTTTTCAGAGCAAGAGATTACGCGCAGACCAAA
ACGATCTCAAGAAGATCATCTTATTAAGGGGTCTGACGCTCAGTGGAACGAAAAATCAATC
TAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCTTAGAAAAACTCATCGAGCATCAA
ATGAAACTGCAATTTATTCATATCAGGATTATCAATACCATATTTTTGAAAAAGCCGTTTCT

GTAATGAAGGAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGT
CTGCGATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAG
GTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATAGCAAAAGCTT
ATGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTC
GCATCAACCAAACCGTTATTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCG
CTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGC
GCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTCC
CGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGATGG
TCGGAAGAGGCATAAATTCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACATCATT
GGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGGCGCATCGGGCTTCCCATACAAT
CGATAGATTGTCGCACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCCATATAAAT
CAGCATCCATGTTGGAATTTAATCGCGGCCTGGAGCAAGACGTTTCCCGTTGAATATGGCT
CATAACACCCCTTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATAT
TTTTATCTTGTGCAATGTACATCAGAGATTTTGAGACACAACCAATTATTGAAGGCCTCCC
TAACGGGGGGCCTTTTTTTGAACGATCGTTGGCTGTGTTGACAATTAATCATCGGCTCGTA
TAATGTGTGGAATTGTGAGCGCTCACAATTGGGTCATGACTGGGACACGCCAGTCATGAG
AATACAGACATAAAGCAAGCCAGAGATTAACGAAGCCAATTATTGAACACCCTTCGGGGT
GTTTTTTTGTTTCTGGTCTACCTACTCCACCGTTGGCTTTTTTCCCTATCAGTGATAGAGAT
TGACATCCCTATCAGTGATAGAGATAATGAGCACGGGCGTTAATCTCTGGCTTGCTTTATG
TCTGTAAACAGAGGAGATACAGAATGAAAGCAAGCAACCTGGCGGCAGCGCAAAAGATGC
GTAAAGCACGTACCCCGAGCCGTAGCAGCATTGGTAGCCTGCGTAGTCCGCATACCCATA
AAGCAATTCTGACCAGCACCATTGAAATCCTGAAAGAATGTGGTTATAGCGGTCTGAGCAT
TGAAAGCGTTGCACGTCGTGCCGGTGCAAGCAAACCGACCATTTATCGTTGGTGGACCAA
TAAAGCAGCACTGATTGCCGAAGTGTATGAAAATGAAAGCGAACAGGTGCGTAAATTTCC
GGATCTGGGTAGCTTTAAAGCCGATCTGGATTTTCTGCTGCGTAATCTGTGGAAAGTTTGG
CGTGAAACCATTTGTGGTGAAGCATTTCGTTGTGTTATTGCAGAAGCACAGCTGGACCCT
GCAACCCTGACCCAGCTGAAAGATCAGTTTATGGAACGTCGTCGTGAGATGCCGAAAAAA
CTGGTTGAAAATGCCATTAGCAATGGTGAACTGCCGAAAGATACCAATCGTGAACTGCTG
CTGGATATGATTTTTGGTTTTTGTTGGTATCGCCTGCTGACCGAACAGCTGACCGTTGAAC
AGGATATTGAAGAATTTACCTTCCTGCTGATTAATGGTGTTTGTCCGGGTACACAGCGTTA
AGGAAACACAGAAAAAAGCCCGCACCTGACAGTGCGGGCTTTTTTTTTCGACCAAAGGTG
TCAACGTTCGACGTACGGTGGAATCTGATTCGTTACCAATTGACATGATACGAAACGTACC
GTATCGTTAAGGTCAAGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGAC
GAAACAGCCTCTACAAATAATTTTGTTTAATACTAGAGAAAGAGGGGAAATACTAGATGGT
GAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCG
ACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGC
AAGCTGACCCTGAAGTTCATCTGCACCACAGGCAAGCTGCCCGTGCCCTGGCCCACCCTC
GTGACCACCTTCGGCTACGGCCTGCAATGCTTCGCCCGCTACCCCGACCACATGAAGCTG

CACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCA
AGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTG
AACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAG
CTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGC
ATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGAC
CACTACCAGCAGAACACCCCAATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTAC
CTTAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTG
CTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAACTC
GGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTTTTCGTTTTGGTCCAGT
TTACGGCTAGCTCAGTCCTAGGTATTATGCTAGCCTGAAGCTGTCACCGGATGTGCTTTCC
GGTCTGATGAGTCCGTGAGGACGAAACAGCCTCTACAAATAATTTTGTTTAATACTAGAGA
AAGAGGGGAAATACTAGATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCC
ATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGG
CGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACAGGCAAGCT
GCCCGTGCCCTGGCCCACCCTCGTGACCACCTTCGGCTACGGCCTGCAATGCTTCGCCCG
CTACCCCGACCACATGAAGCTGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGT
CCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAA
GTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA
CGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCAT
GGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGA
CGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCAATCGGCGACGGCCCCGT
GCTGCTGCCCGACAACCACTACCTTAGCTACCAGTCCGCCCTGAGCAAAGACCCCAACGA
GAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCAT
GGACGAGCTGTACAAGTAACTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGG
GCCTTTTTTCGTTTTGGTCCAATGGCGGCGCGCCATCGAATGGCGCAAAACCTTTCGCGGT
ATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATATGAAACCAGTAACG
TTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACC
AGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTG
AATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGC
GTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCT
CGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGA
AGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCATTAA
CTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGC
GTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAGGAC
GGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTA
GCGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTC
ACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGT
TTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCA

ACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTG
CGGATATCTCGGTAGTGGGATACGACGATACCGAAGATAGCTCATGTTATATCCCGCCGT
TAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGC
AACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCAGTCTCACTGGTGAAAA
GAAAAACCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCAT
TAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGATAATCCAGGAGGA
AAAAAATGTCCAGATTAGATAAAAGTAAAGTGATTAACAGCGCATTAGAGCTGCTTAATGA
GGTCGGAATCGAAGGTTTAACAACCCGTAAACTCGCCCAGAAGCTAGGTGTAGAGCAGCC
TACATTGTATTGGCATGTAAAAAATAAGCGGGCTTTGCTCGACGCCTTAGCCATTGAGATG
TTAGATAGGCACCATACTCACTTTTGCCCTTTAGAAGGGGAAAGCTGGCAAGATTTTTTAC
GTAATAACGCTAAAAGTTTTAGATGTGCTTTACTAAGTCATCGCGATGGAGCAAAAGTACA
TTTAGGTACACGGCCTACAGAAAAACAGTATGAAACTCTCGAAAATCAATTAGCCTTTTTA
TGCCAACAAGGTTTTTCACTAGAGAATGCATTATATGCACTCAGCGCTGTGGGGCATTTTA
CTTTAGGTTGCGTATTGGAAGATCAAGAGCATCAAGTCGCTAAAGAAGAAAGGGAAACAC
CTACTACTGATAGTATGCCGCCATTATTACGACAAGCTATCGAATTATTTGATCACCAAGG
TGCAGAGCCAGCCTTCTTATTCGGCCTTGAATTGATCATATGCGGATTAGAAAAACAACTT
AAATGTGAAAGTGGGTCCTAATAATTGGTAACGAATCAGACAATTGACGGCTCGAGGGAG
TAGCATAGGGTTTGCAGAATCCCTGCTTCGTCCATTTGACAGGCACATTATGCATCGATGA
TAAGCTGTCAAACATGAGCAGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGC
GCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCT
CGCCACTTCGGGCTCATGAGCAAATATTTTATCTGAGGT--3’

A.1.5 pVB005

5’--CTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTC
ATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAG
ATAGGGTTGAGTGGCCGCTACAGGGCGCTCCCATTCGCCATTCAGGCTGCGCAACTGTTG
GGAAGGGCGTTTCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGT
GCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACG
ACGGCCAGTGAGCGCGACGTAATACGACTCACTATAGGGCGAATTGGCGGAAGGCCGTCA
AGGCCACGTGTCTTGTCCAGAGCTCTAATACGACTCACTATAGGGGGGTCATGACTGGGA
CACGCCAGTCATGAGAATACAGACATAAAGCAAGCCAGAGATTAACGAAGCTAGCATAAC
CCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGGGTACCTGGAGCACAAGACTGG
CCTCATGGGCCTTCCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC
ATTAACATGGTCATAGCTGTTTCCTTGCGTATTGGGCGCTCTCCGCTTCCTCGCTCACTGA
CTCGCTGCGCTCGGTCGTTCGGGTAAAGCCTGGGGTGCCTAATGAGCAAAAGGCCAGCAA
AAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCT
GACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA
AGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCG

CTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCAC
GCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAAC
CCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGT
AAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTA
TGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAAC
AGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCT
TGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTA
CGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCA
GTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACC
TAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTG
GTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGT
TCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCAT
CTGGCCCCAGTGCTGCAATGATACCGCGAGAACCACGCTCACCGGCTCCAGATTTATCAG
CAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCT
CCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT
GCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGC
TTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAA
AAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTAT
CACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTT
TTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGT
TGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTG
CTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGAT
CCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAG
CGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGA
CACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGT
TATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTC
CGCGCACATTTCCCCGAAAAGTGCCAC--3’

A.2 Part sequences

Table A.1: Sequences of genetic parts used

Part Name	Type	DNA Sequence (written 5' to 3')
P_{tac}	Promoter [100]	AACGATCGTTGGCTGTGTTGACAATTAATCATCGGCTCGTATAATGT GTGGAATTGTGAGCGCTCACAATT
P_{tet}	Promoter [100]	TACTCCACCGTTGGCTTTTTTCCCTATCAGTGATAGAGATTGACATC CCTATCAGTGATAGAGATAATGAGCAC
P_{phlF}	Promoter [100]	TCTGATTCGTTACCAATTGACATGATACGAAACGTACCGTATCGTTA AGGT
P_{T7}	Promoter [55]	TAATACGACTCACTATAGGG
$sRNA_{20}$	sRNA [55]	GGGTCATGACTGGGACACGCCAGTCATGAGAATACAGACATAAAGCA AGCCAGAGATTAACGAAG
THS_{20}	Toehold switch [55]	GGGCGTTAATCTCTGGCTTGCTTTATGTCTGTAAACAGAGGAGATAC AGAATGAAAGCAAGCAACCTGGCGGCAGCGCAAAAGATGCGTAAA
RiboJ	Insulator [83]	AGCTGTCACCGGATGTGCTTTCCGGTCTGATGAGTCCGTGAGGACGA AACAGCCTCTACAAATAATTTTGTTTAA
yfp	Gene [124]	GTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGT CGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCG AGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATC TGCACCACAGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCAC CTTCGGCTACGGCCTGCAATGCTTCGCCCGCTACCCCGACCACATGA AGCTGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAG GAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGC CGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGA AGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTG GAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCA GAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGG ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCAATC GGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTTAGCTACCA GTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCC TGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAG CTGTACAAGTAA
phlF	Gene [100]	GCACGTACCCCGAGCCGTAGCAGCATTGGTAGCCTGCGTAGTCCGCA TACCCATAAAGCAATTCTGACCAGCACCATTGAAATCCTGAAAGAAT GTGGTTATAGCGGTCTGAGCATTGAAAGCGTTGCACGTCGTGCCGGT GCAAGCAAACCGACCATTTATCGTTGGTGGACCAATAAAGCAGCACT GATTGCCGAAGTGTATGAAAATGAAAGCGAACAGGTGCGTAAATTTC CGGATCTGGGTAGCTTTAAAGCCGATCTGGATTTTCTGCTGCGTAAT CTGTGGAAAGTTTGGCGTGAAACCATTTGTGGTGAAGCATTTCGTTG TGTTATTGCAGAAGCACAGCTGGACCCTGCAACCCTGACCCAGCTGA AAGATCAGTTTATGGAACGTCGTCGTGAGATGCCGAAAAAACTGGTT GAAAATGCCATTAGCAATGGTGAACTGCCGAAAGATACCAATCGTGA ACTGCTGCTGGATATGATTTTTGGTTTTTGTTGGTATCGCCTGCTGA CCGAACAGCTGACCGTTGAACAGGATATTGAAGAATTTACCTTCCTG CTGATTAATGGTGTTTGTCCGGGTACACAGCGTTAA
L3S3P11	Terminator [32]	CCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTACC
L3S2P21	Terminator [32]	CTCGGTACCAAATTCCAGAAAAGAGGCCTCCCGAAAGGGGGGCCTTT TTTCGTTTTGGTCC
ECK120033737	Terminator [32]	GGAAACACAGAAAAAAGCCCGCACCTGACAGTGCGGGCTTTTTTTTT CGACCAAAGG
P_{T7}	Terminator [55]	TAATACGACTCACTATAGGG

A.3 Figures

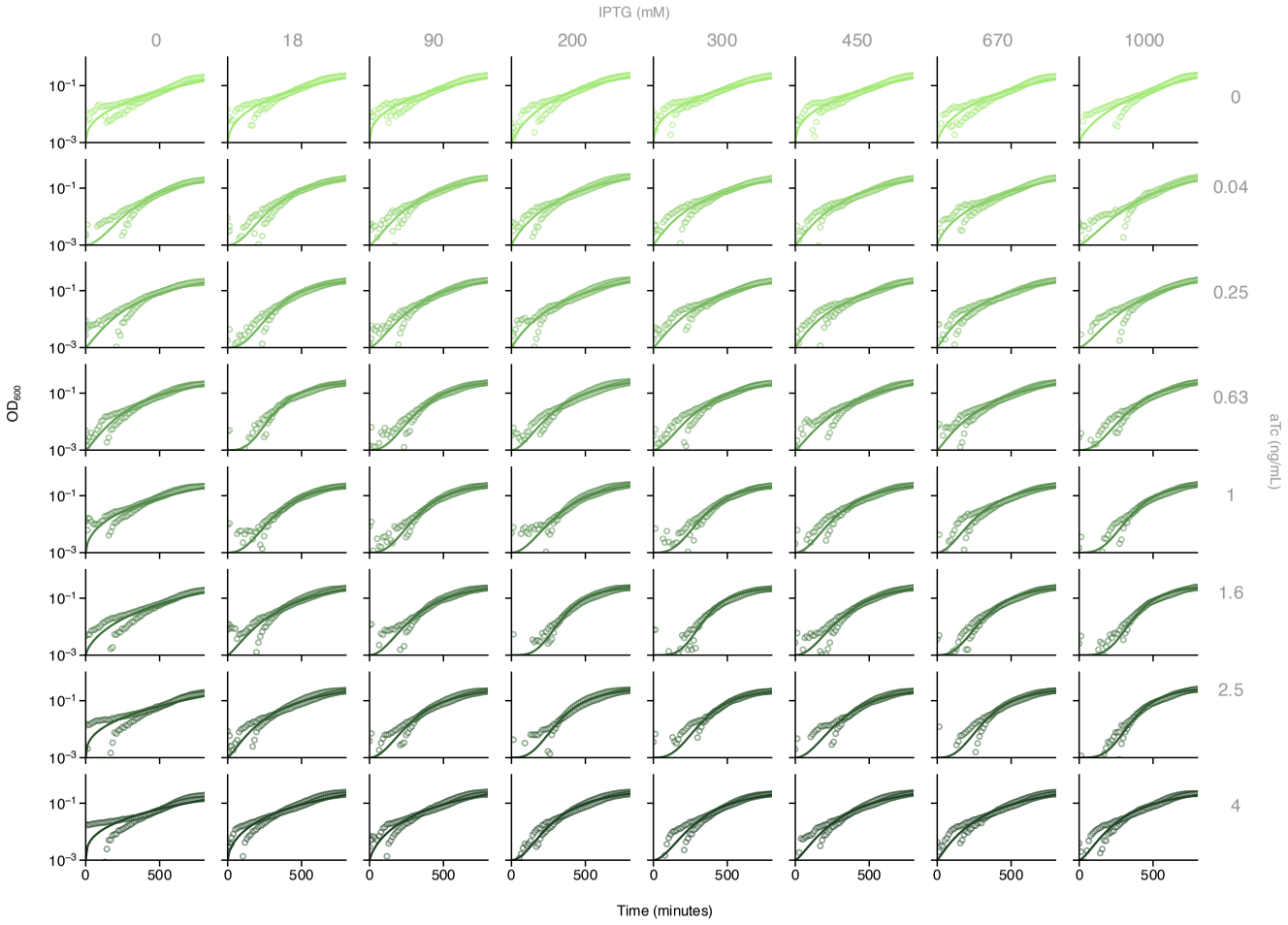


Figure A.1: Growth of cells containing a tuneable expression system. (**A**) Growth curves for populations of cells harbouring the tuneable expression system (TES) grown in inducers aTc and IPTG. All points denote OD_{600} for one of two biological replicates. Solid lines denote growth curve fit to a mathematical model [35].

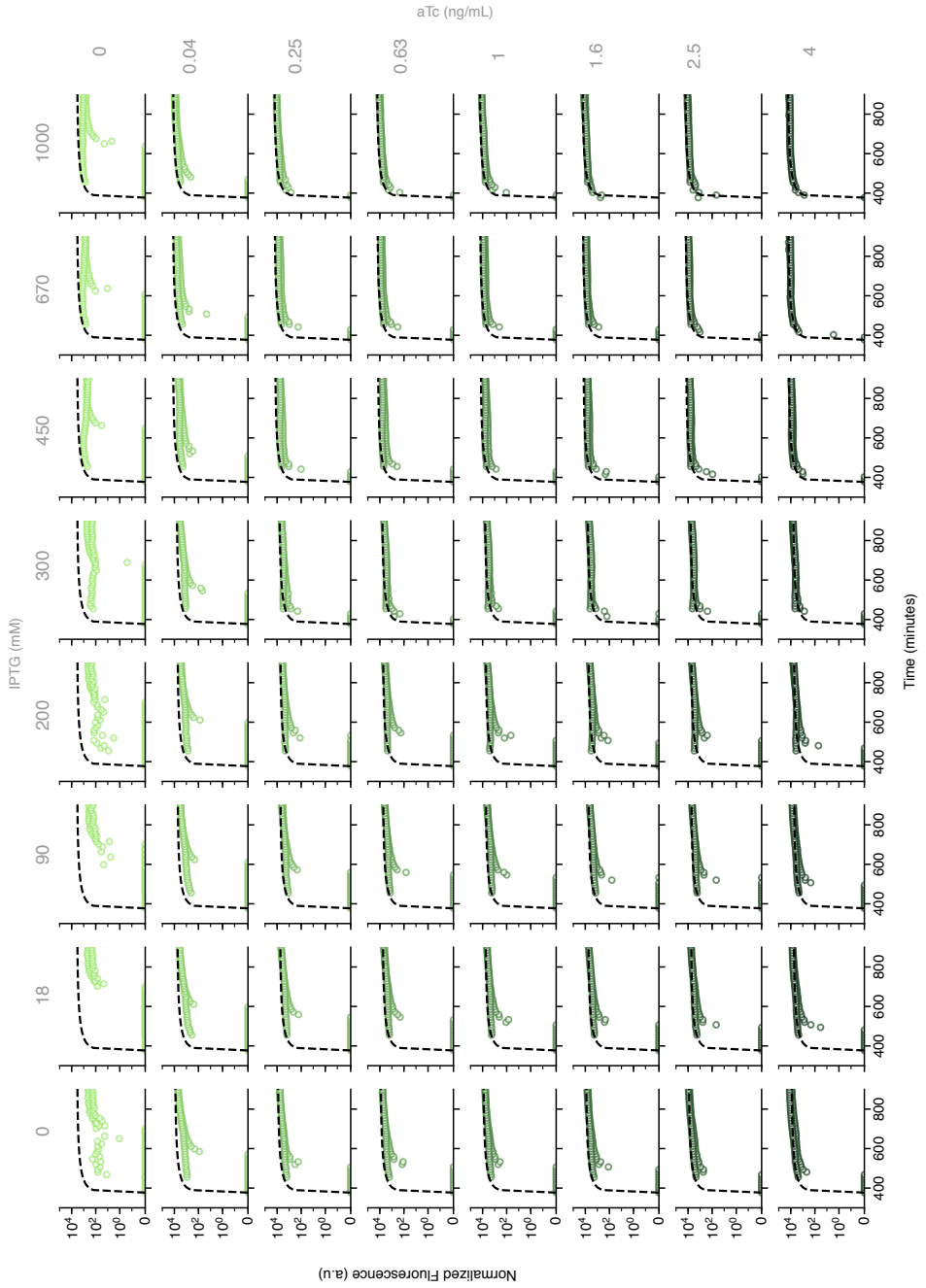


Figure A.2: Change in the rate protein production over time. Points are for single biological replicates. They represent the average concentration of yellow fluorescent protein (YFP), in arbitrary units (a.u) for a bulk of cells. Change in YFP concentration with time is shown for cells grown in varying concentrations of aTc, which increases the input promoter activity and IPTG, which increases the tuner promoter activity. Dashed black line shows the output from the model, **Eq. (4.10 - 4.13)**, fit to the data.

Bibliography

[1] M. ALVAREZ, A. GUZMÁN, AND M. ELÍAS, *Experimental visualization of mixing pathologies in laminar stirred tank bioreactors*, Chemical Engineering Science, 60 (2005), pp. 2449–2457.

[2] J. C. ANDERSON, C. A. VOIGT, AND A. P. ARKIN, *Environmental signal integration by a modular AND gate*, Molecular Systems Biology, 3 (2007), p. 133.

[3] G. BAR-NAHUM AND E. NUDLER, *Isolation and Characterization of σ70-Retaining Transcription Elongation Complexes from Escherichia coli*, p. 9.

[4] V. BARTOLI, G. A. MEAKER, M. DI BERNARDO, AND T. E. GOROCHOWSKI, *Tunable genetic devices through simultaneous control of transcription and translation*, Nature Communications, 11 (2020), p. 2095.

[5] A. BAUDRIMONT, V. JAQUET, S. WALLERICH, S. VOEGELI, AND A. BECSKEI, *Contribution of RNA Degradation to Intrinsic and Extrinsic Noise in Gene Expression*, Cell Reports, 26 (2019), pp. 3752–3761.

[6] J. BEAL, *Signal-to-Noise Ratio Measures Efficacy of Biological Computing Devices and Circuits*, Frontiers in Bioengineering and Biotechnology, 3 (2015).

[7] J. A. BERNSTEIN, A. B. KHODURSKY, P.-H. LIN, S. LIN-CHAO, AND S. N. COHEN, *Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays*, Proceedings of the National Academy of Sciences, 99 (2002), pp. 9697–9702.

[8] M. BHATE, K. MOLNAR, M. GOULIAN, AND W. DEGRADO, *Signal Transduction in Histidine Kinases: Insights from New Structures*, Structure, 23 (2015), pp. 981–994.

[9] S. BIRNBAUM AND J. E. BAILEY, *Plasmid presence changes the relative levels of many host cell proteins and ribosome components in recombinantEscherichia coli*, Biotechnology and Bioengineering, 37 (1991), pp. 736–745.

[10] S. BORUKHOV AND E. NUDLER, *RNA polymerase: the vehicle of transcription*, Trends in Microbiology, 16 (2008), pp. 126–134.

[11] S. BRANTL AND E. G. H. WAGNER, *Antisense RNA-mediated transcriptional attenuation: an in vitro study of plasmid pT181: Antisense RNA mediates transcriptional attenuation*, Molecular Microbiology, 35 (2002), pp. 1469–1482.

[12] E. BRODWIN, *A startup founded by 2 college friends is turning mushrooms into walls and it's already doing deals with major companies like Dell*.
https://www.businessinsider.com/ecovative-turns-mushrooms-into-packaging-ikea-dell-2016-8, Aug. 2016.

[13] J. A. N. BROPHY AND C. A. VOIGT, *Principles of genetic circuit design*, Nature Methods, 11 (2014), pp. 508–520.

[14] R. R. BURGESS AND L. ANTHONY, *How sigma docks to RNA polymerase and what sigma does*, Current Opinion in Microbiology, 4 (2001), pp. 126–131.

[15] J. BUSSGANG AND O. HULL, *C16 Biosciences: Lab-Grown Palm Oil.*, Harvard Business School Case, 820-908 (2019).

[16] J. M. CALLURA, C. R. CANTOR, AND J. J. COLLINS, *Genetic switchboard for synthetic biology applications*, Proceedings of the National Academy of Sciences, 109 (2012), pp. 5850–5855.

[17] J. M. CALLURA, D. J. DWYER, F. J. ISAACS, C. R. CANTOR, AND J. J. COLLINS, *Tracking, tuning, and terminating microbial physiology using synthetic riboregulators*, Proceedings of the National Academy of Sciences, 107 (2010), pp. 15898–15903.

[18] G. CAMBRAY, J. C. GUIMARAES, AND A. P. ARKIN, *Evaluation of 244,000 synthetic sequences reveals design principles to optimize translation in Escherichia coli*, Nature Biotechnology, 36 (2018), pp. 1005–1015.

[19] B. CANTON, A. LABNO, AND D. ENDY, *Refinement and standardization of synthetic biological parts and devices*, Nature Biotechnology, 26 (2008), pp. 787–793.

[20] P. CARBONELL, A. J. JERVIS, C. J. ROBINSON, C. YAN, M. DUNSTAN, N. SWAINSTON, M. VINAIXA, K. A. HOLLYWOOD, A. CURRIN, N. J. W. RATTRAY, S. TAYLOR, R. SPIESS, R. SUNG, A. R. WILLIAMS, D. FELLOWS, N. J. STANFORD, P. MULHERIN, R. LE FEUVRE, P. BARRAN, R. GOODACRE, N. J. TURNER, C. GOBLE, G. G. CHEN, D. B. KELL, J. MICKLEFIELD, R. BREITLING, E. TAKANO, J.-L. FAULON, AND N. S. SCRUTTON, *An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals*, Communications Biology, 1 (2018), p. 66.

[21] S. CARDINALE AND A. P. ARKIN, *Contextualizing context for synthetic biology - identifying causes of failure of synthetic biological systems*, Biotechnology Journal, 7 (2012), pp. 856–866.

[22] T. A. CARRIER AND J. D. KEASLING, *Engineering mRNA Stability in E. coli by the Addition of Synthetic Hairpins Using a 5 Cassette System*, Biotechnology and Bioengineering, 55 (1997), p. 4.

[23] C. CASE, E. SIMONS, AND R. SIMONS, *The IS10 transposase mRNA is destabilized during antisense RNA control.*, The EMBO Journal, 9 (1990), pp. 1259–1266.

[24] I. CASES AND V. DE LORENZO, *Promoters in the environment: transcriptional regulation in its natural context*, Nature Reviews Microbiology, 3 (2005), pp. 105–118.

[25] S. M. CASTILLO-HAIR, J. T. SEXTON, B. P. LANDRY, E. J. OLSON, O. A. IGOSHIN, AND J. J. TABOR, *FlowCal: A User-Friendly, Open Source Software Tool for Automatically Converting Flow Cytometry Data from Arbitrary to Calibrated Units*, ACS Synthetic Biology, 5 (2016), pp. 774–780.

[26] F. CERONI, R. ALGAR, G.-B. STAN, AND T. ELLIS, *Quantifying cellular capacity identifies gene expression designs with reduced burden*, Nature Methods, 12 (2015), pp. 415–418.

[27] F. CERONI, A. BOO, S. FURINI, T. E. GOROCHOWSKI, O. BORKOWSKI, Y. N. LADAK, A. R. AWAN, C. GILBERT, G.-B. STAN, AND T. ELLIS, *Burden-driven feedback control of gene expression*, Nature Methods, 15 (2018), pp. 387–393.

[28] J. CHAPPELL, M. K. TAKAHASHI, AND J. B. LUCKS, *Creating small transcription activating RNAs*, Nature Chemical Biology, 11 (2015), pp. 214–220.

[29] J. CHAPPELL, A. WESTBROOK, M. VEROSLOFF, AND J. B. LUCKS, *Computational design of small transcription activating RNAs for versatile and dynamic gene regulation*, Nature Communications, 8 (2017), p. 1051.

[30] C. CHEN, W. WANG, Z. WANG, F. WEI, AND X. S. ZHAO, *Influence of secondary structure on kinetics and reaction mechanism of DNA hybridization*, Nucleic Acids Research, 35 (2007), pp. 2875–2884.

[31] H. CHEN, M. BJERKNES, R. KUMAR, AND E. JAY, *Determination of the optimal aligned spacing between the Shine â Dalgarno sequence and the translation initiation codon of Escherichia coli m RNAs*, Nucleic Acids Research, 22 (1994), pp. 4953–4957.

[32] Y.-J. CHEN, P. LIU, A. A. K. NIELSEN, J. A. N. BROPHY, K. CLANCY, T. PETERSON, AND C. A. VOIGT, *Characterization of 582 natural and synthetic terminators and quantification of their design constraints*, Nature Methods, 10 (2013), pp. 659–664.

[33] A. CHURSOV, S. J. KOPETZKY, G. BOCHAROV, D. FRISHMAN, AND A. SHNEIDER, *RNAtips: analysis of temperature-induced changes of RNA secondary structure*, Nucleic Acids Research, 41 (2013), pp. W486–W491.

[34] K. P. CLIFTON, E. M. JONES, S. PAUDEL, J. P. MARKEN, C. E. MONETTE, A. D. HALLERAN, L. EPP, AND M. S. SAHA, *The Genetic Insulator RiboJ Increases Expression of Insulated Genes*, preprint, Synthetic Biology, May 2018.

[35] M. CORRADINI AND M. PELEG, *Estimating non-isothermal bacterial growth in foods from isothermal experimental data*, Journal of Applied Microbiology, 99 (2005), pp. 187–200.

[36] F. DELVIGNE, Q. ZUNE, A. R. LARA, W. AL-SOUD, AND S. J. SØRENSEN, *Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity*, Trends in Biotechnology, 32 (2014), pp. 608–616.

[37] U. DEUSCHLE, R. GENTZ, AND H. BUJARD, *lac Repressor blocks transcribing RNA polymerase and terminates transcription.*, Proceedings of the National Academy of Sciences, 83 (1986), pp. 4134–4137.

[38] R. DICKSON, J. ABELSON, W. BARNES, AND W. REZNIKOFF, *Genetic regulation: the Lac control region*, Science, 187 (1975), pp. 27–35.

[39] R. M. DIRKS, J. S. BOIS, J. M. SCHAEFFER, E. WINFREE, AND N. A. PIERCE, *Thermodynamic Analysis of Interacting Nucleic Acid Strands*, SIAM Review, 49 (2007), pp. 65–88.

[40] D. DUANMU, C. BACHY, S. SUDEK, C.-H. WONG, V. JIMENEZ, N. C. ROCKWELL, S. S. MARTIN, C. Y. NGAN, E. N. REISTETTER, M. J. VAN BAREN, D. C. PRICE, C.-L. WEI, A. REYES-PRIETO, J. C. LAGARIAS, AND A. Z. WORDEN, *Marine algae and land plants share conserved phytochrome signaling systems*, Proceedings of the National Academy of Sciences, 111 (2014), pp. 15827–15832.

[41] M. B. ELOWITZ, *Stochastic Gene Expression in a Single Cell*, Science, 297 (2002), pp. 1183–1186.

[42] S.-O. ENFORS, M. JAHIC, A. ROZKOV, B. XU, M. HECKER, B. JÜRGEN, E. KRÜGER, T. SCHWEDER, G. HAMER, D. O'BEIRNE, N. NOISOMMIT-RIZZI, M. REUSS, L. BOONE, C. HEWITT, C. MCFARLANE, A. NIENOW, T. KOVACS, C. TRÄGÅRDH, L. FUCHS, J. REVSTEDT, P. FRIBERG, B. HJERTAGER, G. BLOMSTEN, H. SKOGMAN, S. HJORT, F. HOEKS, H.-Y. LIN, P. NEUBAUER, R. VAN DER LANS, K. LUYBEN, P. VRABEL, AND MANELIUS, *Physiological responses to mixing in large scale bioreactors*, Journal of Biotechnology, 85 (2001), pp. 175–185.

[43] A. ESPAH BORUJENI AND H. M. SALIS, *Translation Initiation is Controlled by RNA Folding Kinetics via a Ribosome Drafting Mechanism*, Journal of the American Chemical Society, 138 (2016), pp. 7016–7023.

[44] X.-J. FENG, S. HOOSHANGI, D. CHEN, G. LI, R. WEISS, AND H. RABITZ, *Optimizing Genetic Circuits by Global Sensitivity Analysis*, Biophysical Journal, 87 (2004), pp. 2195–2202.

[45] C. FLAMM, W. FONTANA, I. L. HOFACKER, AND P. SCHUSTER, *RNA folding at elementary step resolution*, RNA, 6 (2000), pp. 325–338.

[46] T. FRANCH, M. PETERSEN, E. H. WAGNER, J. P. JACOBSEN, AND K. GERDES, *Antisense RNA regulation in prokaryotes: rapid RNA/RNA interaction facilitated by a general U-turn loop structure*, Journal of Molecular Biology, 294 (1999), pp. 1115–1125.

[47] H. GEST, *The discovery of microorganisms by Robert Hooke and Antoni van Leeuwenhoek, Fellows of The Royal Society*, Notes and Records of the Royal Society of London, 58 (2004), pp. 187–201.

[48] A. GHODASARA AND C. A. VOIGT, *Balancing gene expression without library construction via a reusable sRNA pool*, Nucleic Acids Research, 45 (2017), pp. 8116–8127.

[49] W. GILBERT AND A. MAXAM, *The Nucleotide Sequence of the lac Operator*, Proceedings of the National Academy of Sciences, 70 (1973), pp. 3581–3584.

[50] T. E. GOROCHOWSKI, I. CHELYSHEVA, M. ERIKSEN, P. NAIR, S. PEDERSEN, AND Z. IGNATOVA, *Absolute quantification of translational regulation and burden using combined sequencing approaches*, Molecular Systems Biology, 15 (2019), p. e8719.

[51] T. E. GOROCHOWSKI, A. ESPAH BORUJENI, Y. PARK, A. A. NIELSEN, J. ZHANG, B. S. DER, D. B. GORDON, AND C. A. VOIGT, *Genetic circuit characterization and debugging using RNA-seq*, Molecular Systems Biology, 13 (2017), p. 952.

[52] T. E. GOROCHOWSKI, E. VAN DEN BERG, R. KERKMAN, J. A. ROUBOS, AND R. A. L. BOVENBERG, *Using Synthetic Biological Parts and Microbioreactors to Explore the Protein Expression Characteristics of Escherichia coli*, ACS Synthetic Biology, 3 (2014), pp. 129–139.

[53] S. GOTTESMAN, *The Small RNA Regulators of Escherichia coli : Roles and Mechanisms*, Annual Review of Microbiology, 58 (2004), pp. 303–328.

[54] A. A. GREEN, J. KIM, D. MA, P. A. SILVER, J. J. COLLINS, AND P. YIN, *Complex cellular logic computation using ribocomputing devices*, Nature, 548 (2017), pp. 117–121.

[55] A. A. GREEN, P. A. SILVER, J. J. COLLINS, AND P. YIN, *Toehold Switches: De-Novo-Designed Regulators of Gene Expression*, Cell, 159 (2014), pp. 925–939.

[56] T. M. GRUBER AND C. A. GROSS, *Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space*, Annual Review of Microbiology, 57 (2003), pp. 441–466.

[57] A. GYORGY, J. I. JIMÉNEZ, J. YAZBEK, H.-H. HUANG, H. CHUNG, R. WEISS, AND D. DEL VECCHIO, *Isocost Lines Describe the Cellular Economy of Genetic Circuits*, Biophysical Journal, 109 (2015), pp. 639–646.

[58] C. HARINGA, W. TANG, G. WANG, A. T. DESHMUKH, W. A. VAN WINDEN, J. CHU, W. M. VAN GULIK, J. J. HEIJNEN, R. F. MUDDE, AND H. J. NOORMAN, *Computational fluid dynamics simulation of an industrial P. chrysogenum fermentation with a coupled 9-pool metabolic model: Towards rational scale-down and design optimization*, Chemical Engineering Science, 175 (2018), pp. 12–24.

[59] A. HECHT, J. GLASGOW, P. R. JASCHKE, L. A. BAWAZER, M. S. MUNSON, J. R. COCHRAN, D. ENDY, AND M. SALIT, *Measurements of translation initiation from all 64 codons in E. coli*, Nucleic Acids Research, 45 (2017), pp. 3615–3626.

[60] N. HEIDRICH AND S. BRANTL, *Antisense RNA-mediated transcriptional attenuation in plasmid pIP501: the simultaneous interaction between two complementary loop pairs is required for efficient inhibition by the antisense RNA*, Microbiology, 153 (2007), pp. 420–427.

[61] E. HISZCZYŃSKA-SAWICKA AND J. KUR, *Effect of Escherichia coli IHF Mutations on Plasmid p15a Copy Number*, Plasmid, 38 (1997), pp. 174–179.

[62] J. A. HOCH, *Two-component and phosphorelay signal transduction*, Current Opinion in Microbiology, 3 (2000), pp. 165–170.

[63] S. HOOPS, S. SAHLE, R. GAUGES, C. LEE, J. PAHLE, N. SIMUS, M. SINGHAL, L. XU, P. MENDES, AND U. KUMMER, *COPASI–a COmplex PAthway SImulator*, Bioinformatics, 22 (2006), pp. 3067–3074.

[64] L. M. HSU, N. V. VO, C. M. KANE, AND M. J. CHAMBERLIN, *In Vitro Studies of Transcript Initiation by Escherichia coli RNA Polymerase. 1. RNA Chain Initiation, Abortive Initiation, and Promoter Escape at Three Bacteriophage Promoters* [â] , Biochemistry, 42 (2003), pp. 3777–3786.

[65] M. I. HUTCHINGS, H.-J. HONG, AND M. J. BUTTNER, *The vancomycin resistance VanRS two-component signal transduction system of Streptomyces coelicolor*, Molecular Microbiology, 59 (2006), pp. 923–935.

[66] T. N. HUYNH, C. E. NORIEGA, AND V. STEWART, *Missense substitutions reflecting regulatory control of transmitter phosphatase activity in two-component signalling: Non-active-site region K^+ P^- mutants*, Molecular Microbiology, 88 (2013), pp. 459–472.

[67] F. J. ISAACS, D. J. DWYER, AND J. J. COLLINS, *RNA synthetic biology*, Nature Biotechnology, 24 (2006), pp. 545–554.

[68] F. J. ISAACS, D. J. DWYER, C. DING, D. D. PERVOUCHINE, C. R. CANTOR, AND J. J. COLLINS, *Engineered riboregulators enable post-transcriptional control of gene expression*, Nature Biotechnology, 22 (2004), pp. 841–847.

[69] F. JACOB AND J. MONOD, *Genetic regulatory mechanisms in the synthesis of proteins*, p. 39.

[70] M. H. KARAVOLOS, K. WINZER, P. WILLIAMS, AND C. M. A. KHAN, *Pathogen espionage: multiple bacterial adrenergic sensors eavesdrop on host communication systems: Pathogen-host communication*, Molecular Microbiology, 87 (2013), pp. 455–465.

[71] J. KENNEDY, *Particle Swarm Optimization*, in Encyclopedia of Machine Learning and Data Mining, C. Sammut and G. I. Webb, eds., Springer US, Boston, MA, 2017, pp. 967–972.

[72] D. KENNELL AND H. RIEZMAN, *Transcription and translation initiation frequencies of the Escherichia coli lac operon*, Journal of Molecular Biology, 114 (1977), pp. 1–21.

[73] J. KITTLE, R. SIMONS, J. LEE, AND N. KLECKNER, *Insertion sequence IS10 anti-sense pairing initiates by an interaction between the 5â² end of the target RNA and a loop in the anti-sense RNA*, Journal of Molecular Biology, 210 (1989), pp. 561–572.

[74] S. KLUMPP AND T. HWA, *Bacterial growth: global effects on gene expression, growth feedback and proteome partition*, Current Opinion in Biotechnology, 28 (2014), pp. 96–102.

[75] S. KLUMPP, Z. ZHANG, AND T. HWA, *Growth Rate-Dependent Global Effects on Gene Expression in Bacteria*, Cell, 139 (2009), pp. 1366–1375.

[76] N. KOMISSAROVA, J. BECKER, S. SOLTER, M. KIREEVA, AND M. KASHLEV, *Shortening of RNA:DNA Hybrid in the Elongation Complex of RNA Polymerase Is a Prerequisite for Transcription Termination*, Molecular Cell, 10 (2002), pp. 1151–1162.

[77] B. P. LANDRY, R. PALANKI, N. DYULGYAROV, L. A. HARTSOUGH, AND J. J. TABOR, *Phosphatase activity tunes two-component system sensor detection threshold*, Nature Communications, 9 (2018), p. 1433.

[78] K. A. LEAMY, S. M. ASSMANN, D. H. MATHEWS, AND P. C. BEVILACQUA, *Bridging the gap between in vitro and in vivo RNA folding*, Quarterly Reviews of Biophysics, 49 (2016), p. e10.

[79] R. A. LEASE AND M. BELFORT, *A trans-acting RNA as a control switch in Escherichia coli: DsrA modulates function by forming alternative structures*, Proceedings of the National Academy of Sciences, 97 (2000), pp. 9919–9924.

[80] S. Y. LEE AND H. U. KIM, *Systems strategies for developing industrial microbial strains*, Nature Biotechnology, 33 (2015), pp. 1061–1072.

[81] S. T. E. A. LIANG, *Decay of rplN and lacZ mRNA in Escherichia coli*, (1999), p. 18.

[82] C. S. K. LIN, L. A. PFALTZGRAFF, L. HERRERO-DAVILA, E. B. MUBOFU, S. ABDERRAHIM, J. H. CLARK, A. A. KOUTINAS, N. KOPSAHELIS, K. STAMATELATOU, F. DICKSON, S. THANKAPPAN, Z. MOHAMED, R. BROCKLESBY, AND R. LUQUE, *Food waste as a valuable resource for the production of chemicals, materials and fuels. Current situation and global perspective*, Energy & Environmental Science, 6 (2013), p. 426.

[83] C. LOU, B. STANTON, Y.-J. CHEN, B. MUNSKY, AND C. A. VOIGT, *Ribozyme-based insulator parts buffer synthetic circuits from genetic context*, Nature Biotechnology, 30 (2012), pp. 1137–1142.

[84] J. B. LUCKS, L. QI, V. K. MUTALIK, D. WANG, AND A. P. ARKIN, *Versatile RNA-sensing transcriptional regulators for engineering genetic networks*, Proceedings of the National Academy of Sciences, 108 (2011), pp. 8617–8622.

[85] R. LUTZ, *Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements*, Nucleic Acids Research, 25 (1997), pp. 1203–1210.

[86] C. MA AND R. SIMONS, *The IS10 antisense RNA blocks ribosome binding at the transposase translation initiation site.*, The EMBO Journal, 9 (1990), pp. 1267–1274.

[87] C. MADSEN, A. GONI MORENO, Z. PALCHICK, U. P, N. ROEHNER, B. BARTLEY, S. BHATIA, S. BHAKTA, M. BISSELL, K. CLANCY, R. S. COX, T. GOROCHOWSKI, R. GRUNBERG, A. LUNA, J. MCLAUGHLIN, T. NGUYEN, N. LE NOVERE, M. POCOCK, H. SAURO, J. SCOTT-BROWN, J. T. SEXTON, G.-B. STAN, J. J. TABOR, C. A. VOIGT, Z. ZUNDEL, C. MYERS, J. BEAL, AND A. WIPAT, *Synthetic Biology Open Language Visual (SBOL Visual) Version 2.1*, Journal of Integrative Bioinformatics, 16 (2019).

[88] N. MAJDALANI, C. K. VANDERPOOL, AND S. GOTTESMAN, *Bacterial Small RNA Regulators*, Critical Reviews in Biochemistry and Molecular Biology, 40 (2005), pp. 93–113.

[89] O. V. MAKAROVA, E. M. MAKAROV, R. SOUSA, AND M. DREYFUS, *Transcribing of Escherichia coli genes with mutant T7 RNA polymerases: stability of lacZ mRNA inversely correlates with polymerase speed.*, Proceedings of the National Academy of Sciences, 92 (1995), pp. 12250–12254.

[90] M. MARCHISIO AND J. STELLING, *Computational design of synthetic gene circuits with composable parts*, Bioinformatics, 24 (2008), pp. 1903–1910.

[91] P. E. MCGOVERN, J. ZHANG, J. TANG, Z. ZHANG, G. R. HALL, R. A. MOREAU, A. NUNEZ, E. D. BUTRYM, M. P. RICHARDS, C.-S. WANG, G. CHENG, Z. ZHAO, AND C. WANG,

Fermented beverages of pre- and proto-historic China, Proceedings of the National Academy of Sciences, 101 (2004), pp. 17593–17598.

[92] A. J. MEYER, T. H. SEGALL-SHAPIRO, E. GLASSEY, J. ZHANG, AND C. A. VOIGT, *Escherichia coli "Marionette" strains with 12 highly optimized small-molecule sensors*, Nature Chemical Biology, 15 (2019), pp. 196–204.

[93] F. J. MOJICA, C. Dï¿œEZ-VILLASEÏ¿œOR, J. GARCÏ¿œA-MARTÏ¿œNEZ, AND E. SORIA, *Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements*, Journal of Molecular Evolution, 60 (2005), pp. 174–182.

[94] T. S. MOON, C. LOU, A. TAMSIR, B. C. STANTON, AND C. A. VOIGT, *Genetic programs constructed from layered logic gates in single cells*, Nature, 491 (2012), pp. 249–253.

[95] F. MOSER, N. J. BROERS, S. HARTMANS, A. TAMSIR, R. KERKMAN, J. A. ROUBOS, R. BOVENBERG, AND C. A. VOIGT, *Genetic Circuit Performance under Conditions Relevant for Industrial Bioreactors*, ACS Synthetic Biology, 1 (2012), pp. 555–564.

[96] F. MOSER, A. ESPAH BORUJENI, A. N. GHODASARA, E. CAMERON, Y. PARK, AND C. A. VOIGT, *Dynamic control of endogenous metabolism with combinatorial logic circuits*, Molecular Systems Biology, 14 (2018).

[97] K. S. MURAKAMI AND S. A. DARST, *Bacterial RNA polymerases: the wholo story*, Current Opinion in Structural Biology, 13 (2003), pp. 31–39.

[98] V. K. MUTALIK, L. QI, J. C. GUIMARAES, J. B. LUCKS, AND A. P. ARKIN, *Rationally designed families of orthogonal RNA regulators of translation*, Nature Chemical Biology, 8 (2012), pp. 447–454.

[99] K. NATH AND A. L. KOCH, *Protein Degradation in Escherichia coli*, The journal of biological chemistry, 245 (1970).

[100] A. A. K. NIELSEN, B. S. DER, J. SHIN, P. VAIDYANATHAN, V. PARALANOV, E. A. STRYCHALSKI, D. ROSS, D. DENSMORE, AND C. A. VOIGT, *Genetic circuit design automation*, Science, 352 (2016), pp. aac7341–aac7341.

[101] J. NIELSEN AND J. KEASLING, *Engineering Cellular Metabolism*, Cell, 164 (2016), pp. 1185–1197.

[102] E. J. O'BRIEN, J. A. LERMAN, R. L. CHANG, D. R. HYDUKE, AND B. PALSSON, *Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction*, Molecular Systems Biology, 9 (2013), p. 693.

[103] I. OTERO-MURAS AND J. R. BANGA, *Automated Design Framework for Synthetic Biology Exploiting Pareto Optimality*, ACS Synthetic Biology, 6 (2017), pp. 1180–1193.

[104] C. J. PADDON, P. J. WESTFALL, D. J. PITERA, K. BENJAMIN, K. FISHER, D. MCPHEE, M. D. LEAVELL, A. TAI, A. MAIN, D. ENG, D. R. POLICHUK, K. H. TEOH, D. W. REED, T. TREYNOR, J. LENIHAN, H. JIANG, M. FLECK, S. BAJAD, G. DANG, D. DENGROVE, D. DIOLA, G. DORIN, K. W. ELLENS, S. FICKES, J. GALAZZO, S. P. GAUCHER, T. GEISTLINGER, R. HENRY, M. HEPP, T. HORNING, T. IQBAL, L. KIZER, B. LIEU, D. MELIS, N. MOSS, R. REGENTIN, S. SECREST, H. TSURUTA, R. VAZQUEZ, L. F. WESTBLADE, L. XU, M. YU, Y. ZHANG, L. ZHAO, J. LIEVENSE, P. S. COVELLO, J. D. KEASLING, K. K. REILING, N. S. RENNINGER, AND J. D. NEWMAN, *High-level semi-synthetic production of the potent antimalarial artemisinin*, Nature, 496 (2013), pp. 528–532.

[105] A. PAI AND L. YOU, *Optimal tuning of bacterial sensing potential*, Molecular Systems Biology, 5 (2009).

[106] M. C. POLITZ, M. F. COPELAND, AND B. F. PFLEGER, *Artificial repressors for controlling gene expression in bacteria*, Chem. Commun., 49 (2013), pp. 4325–4327.

[107] B. PRESTON, *Synthetic biology as red herring*, Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 44 (2013), pp. 649–659.

[108] L. S. QI, M. H. LARSON, L. A. GILBERT, J. A. DOUDNA, J. S. WEISSMAN, A. P. ARKIN, AND W. A. LIM, *Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression*, Cell, 152 (2013), pp. 1173–1183.

[109] Y. QIAN, H.-H. HUANG, J. I. JIMÉNEZ, AND D. DEL VECCHIO, *Resource Competition Shapes the Response of Genetic Circuits*, ACS Synthetic Biology, 6 (2017), pp. 1263–1272.

[110] A. RAJ AND A. VAN OUDENAARDEN, *Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences*, Cell, 135 (2008), pp. 216–226.

[111] V. RAMAKRISHNAN, *Ribosome Structure and the Mechanism of Translation*, Cell, 108 (2002), pp. 557–572.

[112] N. B. REPPAS, J. T. WADE, G. CHURCH, AND K. STRUHL, *The Transition between Transcriptional Initiation and Elongation in E. coli Is Highly Variable and Often Rate Limiting*, Molecular Cell, 24 (2006), pp. 747–757.

[113] D. RONCARATI AND V. SCARLATO, *Regulation of heat-shock genes in bacteria: from signal sensing to gene expression output*, FEMS Microbiology Reviews, 41 (2017), pp. 549–574.

[114] R. M. SAECKER, M. T. RECORD, AND P. L. DEHASETH, *Mechanism of Bacterial Transcription Initiation: RNA Polymerase - Promoter Binding, Isomerization to Initiation-*

Competent Open Complexes, and Initiation of RNA Synthesis, Journal of Molecular Biology, 412 (2011), pp. 754–771.

[115] H. M. SALIS, E. A. MIRSKY, AND C. A. VOIGT, *Automated design of synthetic ribosome binding sites to control protein expression*, Nature Biotechnology, 27 (2009), pp. 946–950.

[116] P. SANCHEZ-VAZQUEZ, C. N. DEWEY, N. KITTEN, W. ROSS, AND R. L. GOURSE, *Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase*, Proceedings of the National Academy of Sciences, 116 (2019), pp. 8310–8319.

[117] R. SANDERS, *Brewing hoppy beer without the hops*. https://news.berkeley.edu/2018/03/20/brewing-hoppy-beer-without-the-hops/, Mar. 2018.

[118] T. M. SCHMEING AND V. RAMAKRISHNAN, *What recent ribosome structures have revealed about the mechanism of translation*, Nature, 461 (2009), pp. 1234–1242.

[119] M. A. SELLITTI, P. A. PAVCO, AND D. A. STEEGE, *lac repressor blocks in vivo transcription of lac control region DNA.*, Proceedings of the National Academy of Sciences, 84 (1987), pp. 3199–3203.

[120] M. E. SHERLOCK, N. SUDARSAN, AND R. R. BREAKER, *Riboswitches for the alarmone ppGpp expand the collection of RNA-based signaling systems*, Proceedings of the National Academy of Sciences, 115 (2018), pp. 6052–6057.

[121] J. D. SIMEN, M. LÖFFLER, G. JÄGER, K. SCHÄFERHOFF, A. FREUND, J. MATTHES, J. MÜLLER, R. TAKORS, AND RECOGNICE-TEAM, *Transcriptional response of Escherichia coli to ammonia and glucose fluctuations*, Microbial Biotechnology, 10 (2017), pp. 858–872.

[122] A. SIMONETTI, S. MARZI, L. JENNER, A. MYASNIKOV, P. ROMBY, G. YUSUPOVA, B. P. KLAHOLZ, AND M. YUSUPOV, *A structural view of translation initiation in bacteria*, Cellular and Molecular Life Sciences, 66 (2009), pp. 423–436.

[123] A. SNIR, D. NADEL, I. GROMAN-YAROSLAVSKI, Y. MELAMED, M. STERNBERG, O. BAR-YOSEF, AND E. WEISS, *The Origin of Cultivation and Proto-Weeds, Long Before Neolithic Farming*, PLOS ONE, 10 (2015), p. e0131422.

[124] B. C. STANTON, A. A. K. NIELSEN, A. TAMSIR, K. CLANCY, T. PETERSON, AND C. A. VOIGT, *Genomic mining of prokaryotic repressors for orthogonal logic gates*, Nature Chemical Biology, 10 (2014), pp. 99–105.

[125] E. J. STEEN, Y. KANG, G. BOKINSKY, Z. HU, A. SCHIRMER, A. MCCLURE, S. B. DEL CARDAYRE, AND J. D. KEASLING, *Microbial production of fatty-acid-derived fuels and chemicals from plant biomass*, Nature, 463 (2010), pp. 559–562.

[126] M. J. SWAIN AND D. H. BALLARD, *Color Indexing*, International Journal of Computer Vision, 7 (1991), pp. 11–32.

[127] M. K. TAKAHASHI AND J. B. LUCKS, *A modular strategy for engineering orthogonal chimeric RNA transcription regulators*, Nucleic Acids Research, 41 (2013), pp. 7577–7588.

[128] A. TAMSIR, J. J. TABOR, AND C. A. VOIGT, *Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'*, Nature, 469 (2011), pp. 212–215.

[129] G. TOWLER AND R. SINNOTT, *Chapter 9 - economic evaluation of projects*, in Chemical Engineering Design (Second Edition), G. Towler and R. Sinnott, eds., Butterworth-Heinemann, Boston, second edition ed., 2013, pp. 389 – 429.

[130] P. VAIDYANATHAN, B. S. DER, S. BHATIA, N. ROEHNER, R. SILVA, C. A. VOIGT, AND D. DENSMORE, *A Framework for Genetic Logic Synthesis*, Proceedings of the IEEE, 103 (2015), pp. 2196–2207.

[131] L. C. VITORINO AND L. A. BESSA, *Technological Microbiology: Development and Applications*, Frontiers in Microbiology, 8 (2017), p. 827.

[132] C. A. VOIGT, *Genetic parts to program bacteria*, Current Opinion in Biotechnology, 17 (2006), pp. 548–557.

[133] P. H. VON HIPPEL, *An Integrated Model of the Transcription Complex in Elongation, Termination, and Editing*, Science, 281 (1998), pp. 660–665.

[134] J. T. WADE, D. C. ROA, D. C. GRAINGER, D. HURD, S. J. W. BUSBY, K. STRUHL, AND E. NUDLER, *Extensive functional overlap between Ï factors in Escherichia coli*, Nature Structural & Molecular Biology, 13 (2006), pp. 806–814.

[135] A. Y. WEISSE, D. A. OYARZÚN, V. DANOS, AND P. S. SWAIN, *Mechanistic links between cellular trade-offs, gene expression, and growth*, Proceedings of the National Academy of Sciences, 112 (2015), pp. E1038–E1047.

[136] K. S. WILSON AND C. R. CONANT, *Determinants of the Stability of Transcription Elongation Complexes: Interactions of the Nascent RNA with the DNA Template and the RNA Polymerase*, Journal of Molecular Biology, (1999), p. 16.

[137] S. XU, J. ZHAN, B. MAN, S. JIANG, W. YUE, S. GAO, C. GUO, H. LIU, Z. LI, J. WANG, AND Y. ZHOU, *Real-time reliable determination of binding kinetics of DNA hybridization using a multi-channel graphene biosensor*, Nature Communications, 8 (2017), p. 14902.

[138] Y. YOKOBAYASHI, R. WEISS, AND F. H. ARNOLD, *Directed evolution of a genetic circuit*, Proceedings of the National Academy of Sciences, 99 (2002), pp. 16587–16591.

[139] S.-H. YOON, J.-E. KIM, S.-H. LEE, H.-M. PARK, M.-S. CHOI, J.-Y. KIM, S.-H. LEE, Y.-C. SHIN, J. D. KEASLING, AND S.-W. KIM, *Engineering the lycopene synthetic pathway in E. coli by comparison of the carotenoid genes of Pantoea agglomerans and Pantoea ananatis*, Applied Microbiology and Biotechnology, 74 (2007), pp. 131–139.

[140] C. YOU, H. OKANO, S. HUI, Z. ZHANG, M. KIM, C. W. GUNDERSON, Y.-P. WANG, P. LENZ, D. YAN, AND T. HWA, *Coordination of bacterial proteome with metabolism by cyclic AMP signalling*, Nature, 500 (2013), pp. 301–306.

[141] G. Z. YUSUPOVA, M. M. YUSUPOV, J. CATE, AND H. F. NOLLER, *The Path of Messenger RNA through the Ribosome*, Cell, 106 (2001), pp. 233–241.

[142] F. ZHANG, J. M. CAROTHERS, AND J. D. KEASLING, *Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids*, Nature Biotechnology, 30 (2012), pp. 354–359.